

**Aus der Radiologischen Universitätsklinik (Department)
Tübingen**

Abteilung Diagnostische und Interventionelle Radiologie

Ärztlicher Direktor: Professor Dr. C. D. Claussen

**Evaluation eines Matrix-Systems zur Klassifikation von
Mikroverkalkungen in Mammographien –
Eine „Multi-Reader“- Studie**

**Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin**

**der Medizinischen Fakultät
der Eberhard-Karls-Universität
zu Tübingen**

**vorgelegt von
Rebecca Sarah Läufler**

aus

Radolfzell am Bodensee

2008

Dekan:

Professor Dr. I. B. Autenrieth

1. Berichterstatter:

Professor Dr. S. Miller

2. Berichterstatter:

Professor Dr. E. Solomayer

Inhaltsverzeichnis

1. Einleitung	1
1.1 Das Mammakarzinom	1
1.1.1 Epidemiologie	1
1.1.2 Stadienabhängige Überlebensraten	3
1.1.3 DCIS und LCIS.....	5
1.2 Die Mammographie	6
1.3 Verkalkungen	9
1.3.1 Verkalkungen in der Mammographie.....	9
1.3.2 Makroverkalkungen	10
1.3.3 Mikroverkalkungen	10
1.4 BI-RADS	12
1.4.1 Allgemeines.....	12
1.4.2 Was sagt BI-RADS über Verkalkungen?	14
1.5 Aufgabenstellung und Zielsetzung	19
2. Material und Methoden	21
2.1 Patienten	21
2.2 Diagnostische Verfahren	22
2.2.1 Mammographie	22
2.2.2 Biopsieverfahren und Histologie	23
2.3 Aufbereitung und Auswertung der Patientendaten	24
2.3.1 Aufbereitung der Daten.....	24
2.3.2 Befundung der Mammographien	25
2.3.3 Das Mikrokalkmatrixsystem	28
2.3.4 Auswertung	28
2.4 Statistische Erhebungen	30
3. Ergebnisse	32
3.1. Histologisch-Pathologische Ergebnisse	32
3.2. Statistische Ergebnisse	35
3.2.1 Methode 1: Auswertung mit dem Matrixschema	35
3.2.1.1 Sensitivität	35
3.2.1.2 Spezifität.....	38

3.2.2 Methode 2: Auswertung nach der subjektiven Beurteilung	40
3.2.2.1 Sensitivität	40
3.2.2.2 Spezifität	43
3.2.3 Methode 3: Auswertung mit der modifizierten Matrix	46
3.2.3.1 Sensitivität	46
3.2.3.2 Spezifität	48
3.2.4 Vergleich der Gesamtsensitivitäten und –spezifitäten	51
3.2.5 Vergleich der Readings	54
3.2.5.1 Reader 1	54
3.2.5.2 Reader 2	55
3.2.5.3 Reader 3	56
3.2.5.4 Reader 4	57
3.2.6 Übereinstimmungen	58
3.2.6.1 Methode 1	58
3.2.6.2 Methode 2	61
3.2.6.3 Methode 3	63
3.2.7 Ergebnisse nach BI-RADS-Kategorie	66
3.2.8 Verteilung der Malignitätswahrscheinlichkeit über die Matrix	70
3.3. Beispielmammographien	72
4. Diskussion	78
4.1 Möglichkeiten und Grenzen von BI-RADS	78
4.2 Diskussion der Ergebnisse	79
4.2.1 Histologisch-Pathologische Ergebnisse	79
4.2.2 Statistische Ergebnisse	82
4.2.2.1 Sensitivitäten und Spezifitäten	82
4.2.2.2 Vergleich der Readings	85
4.2.2.3 Übereinstimmungen	86
4.2.2.4 Kontroverse BI-RADS Kategorie 3	88
4.2.2.5 Ergebnisse nach BI-RADS-Kategorie	89
4.2.2.6 Verteilung der Malignitätswahrscheinlichkeit über die Matrix ...	92
4.3 Abschließende Beurteilung und Ausblick	95
5. Zusammenfassung	98
6. Anhang	100
6.1 Tabellenverzeichnis	100
6.2 Abbildungsverzeichnis	102

Inhaltsverzeichnis

6.3 Abkürzungen	104
6.4 Abstract.....	105
7. Literaturverzeichnis	107
8. Danksagung	118
9. Lebenslauf.....	119

1. Einleitung

1.1 Das Mammakarzinom

1.1.1 Epidemiologie

Das Mammakarzinom ist die häufigste bösartige Neubildung der Frau in Deutschland und auch weltweit. Neben dem oft tragischen persönlichen Schicksal der betroffenen Frau stellt es durch seine zunehmende Inzidenz ein großes gesundheitspolitisches und sozialpolitisches Problem dar [1, 2].

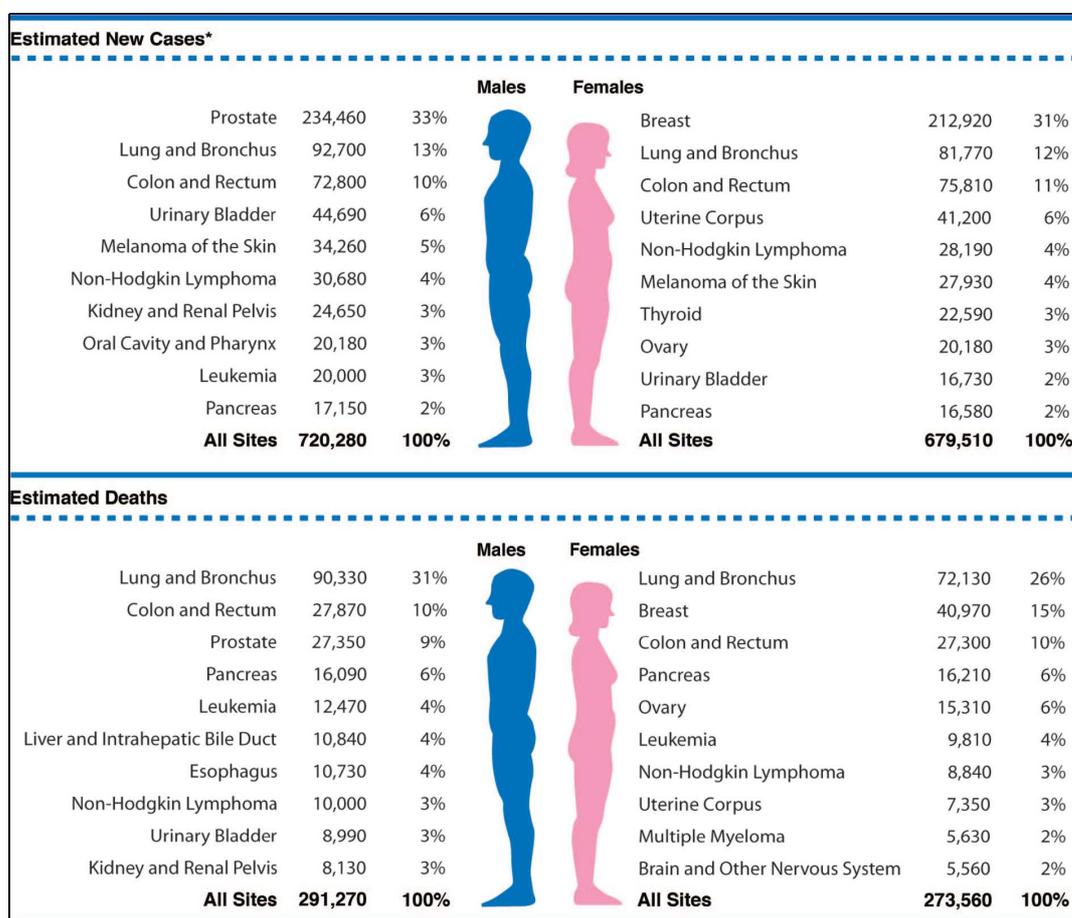


Abb. 1: Die zehn führenden Krebsarten mit den geschätzten neu aufgetretenen Fällen und Todesfällen in den USA 2006, nach Geschlecht getrennt

* Ohne Basal- und Plattenepithelkazinome der Haut und 'In-Situ Carcinoma' (außer Harnblase) Abbildung übernommen aus 'Cancer Statistics 2006' [3]

Das mittlere Lebenszeitrisiko für die Entstehung dieses Tumors beträgt in Deutschland 9,2%. Somit erkrankt jede 11. Frau im Laufe ihres Lebens an

Brustkrebs [4], mit weltweit steigender Tendenz [5-8]. In Deutschland erkranken jährlich etwa 55100 Frauen [9]. Dies entspricht einer Inzidenz von 129,8/100000 Einwohner [9].

Die Wahrscheinlichkeit, an einem Karzinom der Brust zu erkranken nimmt mit dem Alter zu, jedoch ist der Brustkrebs auch die häufigste Krebserkrankung bei jüngeren Frauen. 10% der Frauen sind bei Diagnosestellung jünger als 45 Jahre. Das mittlere Erkrankungsalter liegt bei 62 Jahren [9]. Dies ist im Gegensatz zu vielen anderen Malignomen vergleichsweise früh.

Das Vorliegen von Risikofaktoren erhöht ebenfalls die Wahrscheinlichkeit für das Auftreten eines Malignoms. Neben einem Lebensalter über 50 Jahre, reicht auch eine familiäre Disposition mit Verwandtschaft 1.Grades allein aus, um das Risiko an Brustkrebs zu erkranken, zu verdoppeln. Als weitere gesicherte Faktoren, welche direkt die Häufigkeit einer Erkrankung beeinflussen, sind eine frühe Menarche, eine späte Menopause, Nulliparität, ionisierende Strahlung und Adipositas zu nennen [10]. Die individuelle Vorbelastung durch ein Karzinom der kontralateralen Brust oder einer Karzinomvorstufe im Sinne einer atypischen duktaalen oder lobulären Hyperplasie erhöht ebenfalls das Brustkrebsrisiko [11].

Weiterhin sind bei 5 – 10% aller Mammakarzinome genetische Ursachen im Sinne von Mutationen bei Onkogenen beziehungsweise bei Tumorsuppressorgenen bekannt [12].

In Deutschland führt der Brustkrebs in der Altersgruppe der 39- bis 58-jährigen Frauen weiterhin mit bis zu 15% die Todesursachenstatistik an [13]. Wie in Tabelle 1 dargestellt, gingen in den USA mit 31%, die meisten aller neuen Krebsfälle bei Frauen im Jahr 2006 zulasten des Mammakarzinoms [3].

Betrachtet man jedoch die Entwicklung der Mortalitätsraten des Mammakarzinoms, so findet man in den Jahren von 1990 bis 2000 in den USA einen Rückgang von 25%, in Großbritannien sogar von 30%. Auch in

Deutschland ist seit wenigen Jahren ein Rückgang zu verzeichnen. Die Gründe hierfür liegen vor allem in einer verbesserten Früherkennung, einem erhöhtem Brustkrebsbewusstsein in der Bevölkerung und einer höheren Effektivität von neuen Therapien. In Deutschland ist erst im Jahr 2007 ein bundesweites Screening eingeführt worden. Aus diesem Grund lassen sich die genauen Effekte und Auswirkungen auf die Mortalität erst ab dem Jahr 2017 beobachten. Die epidemiologisch herausragende Bedeutung des Mammakarzinoms hat zur Folge, dass schon geringe Verbesserungen bei der Versorgung einen großen Effekt auf die Gesundheit der Gesamtbevölkerung haben [11].

1.1.2 Stadienabhängige Überlebensraten

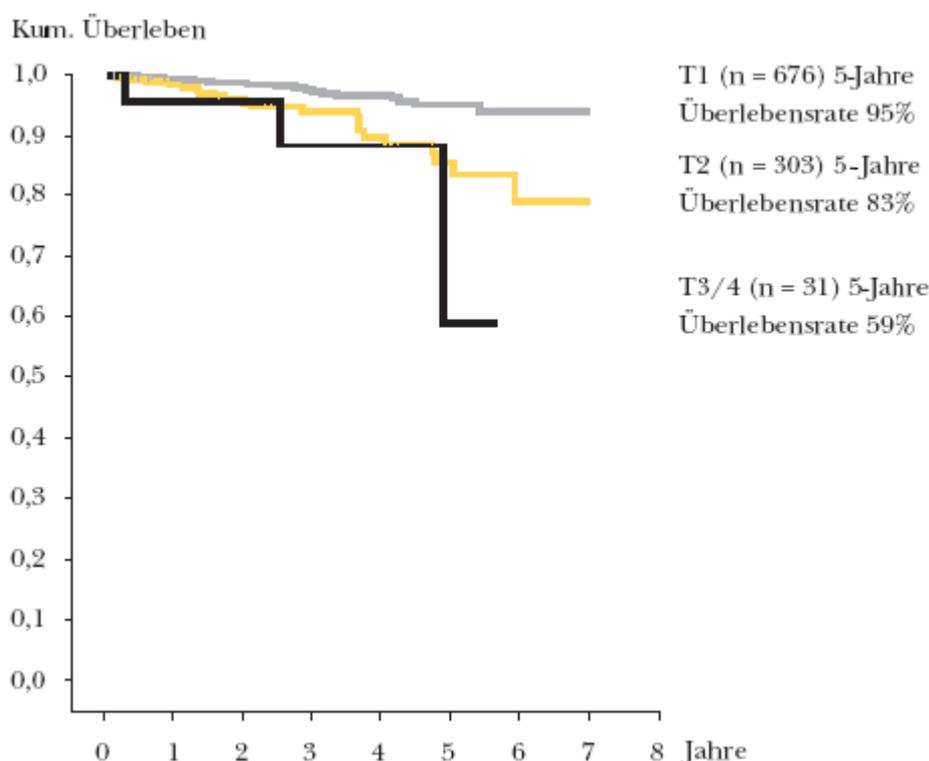


Abb. 2: T-Stadienabhängige Überlebensraten des Mammakarzinoms bei NO, MO. Abbildung übernommen von 'Onkologischer Schwerpunkt', HSK Wiesbaden 2004 [14]

Die Überlebensraten des Mammakarzinoms unterscheiden sich stark in Abhängigkeit vom Stadium der Erkrankung bei Diagnosestellung. Wie in Abbildung 2 ersichtlich beträgt die 5-Jahres-Überlebensrate im Stadium pT1 95%, im Stadium pT3 hingegen nur noch 59%. Beim Vorliegen einer

metastasierten Brustkrebserkrankung überleben hingegen nur noch 18% der Betroffenen. Die mediane 5-Jahres-Überlebensrate des Mammakarzinoms liegt bei 84% [4, 14].

Wie bei der Mehrzahl der malignen Tumorerkrankungen gilt auch das Mammakarzinom als Systemerkrankung, da bei Entdeckung des Primärtumors schon okkulte Fernmetastasen vorliegen können. Insbesondere das Mammakarzinom neigt zu früher hämatogener und lymphogener Metastasierung.

Dabei korreliert die Wahrscheinlichkeit für eine Metastasierung direkt zur Tumorgöße. Durch die deutliche positive Korrelation des Krankheitsstadiums mit der Größe des Primärtumors, ist schließlich auch die Prognose der Tumorerkrankung eng mit der Tumorgöße verknüpft.

Diese Beobachtungen zeigen deutlich, wie wichtig es ist, den Krebs in einem frühen Stadium zu entdecken. Durch eine zuverlässige gynäkologische Vorsorge unter Einbeziehung der Mammadiagnostik ist somit eine deutliche Verbesserung der Heilungs- und Überlebensraten von betroffenen Patientinnen möglich. Da es gegenwärtig keine effektive Methode der Primärprävention gibt ist die Mammographie als Sekundärprävention die zuverlässigste Möglichkeit Brustkrebs in einem frühen Stadium zu diagnostizieren [1, 15, 16]. Neben der Mammographie werden in der Brustdiagnostik auch die Sonographie, MRT-Untersuchungen und die klinische Untersuchung angewandt.

Die Prognose der anlässlich einer Vorsorgeuntersuchung entdeckten Mammakarzinome ist wesentlich besser als die von bereits klinisch manifesten Karzinomen [2]. Das Tumorstadium ist günstiger, Lymphknotenbefall und Fernmetastasierung sind seltener.

1.1.3 DCIS und LCIS

Hinsichtlich der Histologie kann man die nichtinvasiven Karzinome von invasiven Karzinomen unterscheiden. Zum erstgenannten Typus zählen das intraduktale Karzinom und das Paget- Karzinom. Bei den invasiven Karzinomen gibt es eine Vielzahl unterschiedlicher histologischer Entitäten. Das invasiv-duktales Karzinom stellt mit 65 % das mit Abstand häufigste Karzinom dieser Untergruppe dar. Weitere invasive Karzinome sind unter anderem das invasiv-lobuläre Karzinom mit etwa 10%, das invasiv- papilläre Karzinom, das tubuläre Karzinom und das muzinöse Karzinom.

Ein Anteil von 15-20% der malignen Mammakarzinome sind präinvasive Neoplasien, also lobuläre und intraduktale Veränderungen. Diese überschreiten die epitheliale Basalmembran nicht und wachsen somit nicht ins umliegende Brustdrüsenstroma ein. Hierzu zählt man das duktales Carcinoma-in-situ (= DCIS) und das lobuläre Carcinoma-in-situ (= LCIS).

Obwohl die genaue Tumorgenese des Brustkrebses nur teilweise bekannt ist, wird es doch als sehr wahrscheinlich angesehen, dass nahezu alle invasiven Mammakarzinome aus in-situ-Vorstufen entstehen.

Das DCIS ist eine vor der Invasion stehende maligne neoplastische Veränderung, die sich aus normalen Zellen des duktales Lumens über benigne proliferative Vorstufen und atypische Hyperplasien zum Carcinoma-in-situ entwickelt [17]. Nach Klassifikation der World Health Organisation (= WHO) von 2003 kann man die duktales intraepithelialen Neoplasien (DIN) histologisch nach DIN 1A und DIN 1B einteilen. Dann entspricht DIN 1A der duktales Hyperplasie, DIN 1B der atypischen Hyperplasie. Das duktales Carcinoma-in-situ selbst wird in DIN 1C als „low grade“, DIN 2 als „intermediate grade“ und DIN 3 als „high grade“ unterteilt [18, 19].

Das DCIS wird in den meisten Fällen mammographisch diagnostiziert und kann daher durch Screening- Untersuchungen frühzeitig entdeckt werden [20]. Aufgrund eines großflächig eingesetzten Mammographie- Screenings ist die Inzidenz des duktales Carcinoma-in-situ in den letzten 20 Jahren dramatisch

angestiegen [17]. Nach der Studie von Ernster et. al sind 20% aller durch Screening entdeckten Karzinome DCIS [21].

1.2 Die Mammographie

Die Mammographie ist erstmals 1927 durch Kleinschmidt [22] eingeführt. Sie gilt, neben der Palpation als das älteste nichtinvasive Untersuchungsverfahren zum Nachweis von Brustkrebs.

Neben malignen Veränderungen sind auch zahlreiche benigne Veränderungen, darunter Mastopathien, Zysten, Fibroadenome, Papillome und Entzündungen der Brustdrüse durch die Mammographie diagnostizierbar.

Eine herausragende Rolle spielt die Mammographie aber vor allem bei der Früherkennung kleiner und kleinster Karzinome die noch klinisch unauffällig sind. Die Mammographie ist in der Lage okkulte Neoplasien ab einem Größendurchmesser von wenigen Millimetern nachzuweisen.

Zahlreiche Screeningprogramme belegen eine hohe Sensitivität dieser Methode. Insgesamt gesehen reichen in der Literatur die Daten für die Sensitivität von 85-90% bei Frauen über 50 Jahren [23-26]. Die Werte für die Spezifität liegen bei 90-95% [27, 28].

Verschiedene Studien belegen außerdem eine Verbesserung der Mortalitäts- und Morbiditätsraten durch das mammographische Screening [29-31]. Die meisten Experten sehen in der Routinemammographie die effektivste Methode zur Brustkrebsprävention von Frauen über 50 Jahren [32-35].

Standardmäßig wird die Mammographie in 2 Ebenen durchgeführt. Zum einen in der craniocaudalen Projektion (= CC) und zum anderen in der Schräg-Projektion, der medio-lateral-obliquen Projektion (= MLO). Dadurch wird die Fehlerrate beim Befunden der Bilder im Vergleich mit der Einebenen-Mammographie reduziert [36].

Zusatzaufnahmen sind die medio- laterale (= ML) Aufnahme und

Kompressionsvergrößerungen. Kompressionsvergrößerungen senken die Bewegungsunschärfe und den Streustrahlenanteil.

Eine besondere Bedeutung haben die durch die Mammographie darstellbaren Verkalkungen. Mikroverkalkungen sind häufig das einzige radiologische Malignitätszeichen [37-39]. Die Schwierigkeit besteht darin, dass es häufig mammographisch nur schwer möglich ist, den Mikrokalk in benigne und maligne Befunde zu differenzieren [38, 40-42]. Durch Vergrößerungsaufnahmen können die Mikroverkalkungen besser beurteilt werden.

Die Mammographie hat jedoch auch ihre Grenzen. So sind bis zu 10% aller Karzinome, die ausschließlich der Palpation zugänglich sind, mammographisch nicht darstellbar. Darüber hinaus ist bei extrem dichtem Drüsengewebe, besonders bei prämenopausalen Frauen und bei mastopathischem Brustdrüsengewebe die Genauigkeit der radiologischen Mammabeurteilung erniedrigt [43]. Des Weiteren ist auch nach Segmentresektion von Karzinomen, Augmentationsplastiken und Strahlentherapie die Aussagekraft herabgesetzt [44-46].

Ein negatives Mammographie- Screening schließt also ein Malignom der Brust nicht aus. Das ACR (= ‚*American College of Radiology*‘) weist deshalb deutlich daraufhin, dass zur präzisen Beurteilung auch Maßnahmen wie die Tastuntersuchung, die Sonographie und die Biopsie der Brust gehören. Die Kombination dieser Untersuchungen ermöglicht eine Sensitivität bis zu 97% [47].

Das Leistungsniveau variiert stark zwischen den Radiologen die Mammographien befunden. Dies hängt zum einen von der geringen Inzidenz des Brustkrebses ab und zum anderen von der Schwierigkeit verdächtige Areale vom umgebenden Brustdrüsengewebe abzugrenzen und zu identifizieren. Nicht zuletzt mag die ermüdende und wiederholende Natur des Mammographie- Befundens auch ein Teil dazu beitragen [48-50].

Eine anerkannte Strategie um Fehler in der Befundung von Mammographien zu reduzieren ist das ‚double reading‘. Dabei wird die Mammographie von zwei

Radiologen unabhängig befundet [51, 52] .

Ein großer Fortschritt in der Mammographietechnik brachten die achtziger und neunziger Jahre des letzten Jahrhunderts mit der Entwicklung der digitalen Mammographie (Full Field Digital Mammography = FFDM) [53-55]. Sie erlaubt eine bessere Auflösung der Bildkontraste was besonders für dichtes Gewebe von Vorteil ist [1].

Durch die Möglichkeit der Bearbeitung mit moderner Software soll mit der digitalen Mammographie die Genauigkeit bei der Diagnostizierung von Brustkrebs verbessert werden [56]. In der Literatur wird dies jedoch eher kontrovers diskutiert [56-58].

In den Studie von Fischer et. al und Fischmann et. al zeigt sich die digitale Mammographie bei der Detektion von Mikroverkalkungen der Film-Folien Mammographie (Film Screen Mammographie = FSM) klar überlegen [59, 60].

Die Mammographie stellt im Vergleich zu anderen radiologischen Untersuchungen wie beispielsweise einer Röntgenaufnahme des Thorax eine relativ hohe Strahlenexposition für die Patientin dar. Aus diesem Grund wird die strahlungsbedingte Krebsinduktion kontrovers diskutiert.

Durch eine einmalige Mammographieuntersuchung wird das Risiko, an Brustkrebs zu erkranken rechnerisch, von 12% auf 12,0036% erhöht. Jenseits des 40. Lebensjahres sind die Risiken der Mammographie gegenüber dem Nutzen zu vernachlässigen [61, 62]. Bei den unter 40jährigen ist auf Grund des wesentlich geringeren Erkrankungsrisikos, der höheren Strahlensensibilität und des dichten Parenchyms ein Screening nicht sinnvoll [63].

Im Allgemeinen ist das Strahlenrisiko durch die Mammographie als äußerst gering anzusehen [2, 31, 64].

Tab. 1: Spezielle Tumore/Mammographie

Spezielle Tumoren	Mammographie
LCIS	Meist unspezifisch
DCIS	Meist gut erkennbar (bis zu 85% Mikrokalk)
Duktales Karzinom	Meist gut erkennbar, Schwierigkeiten bereiten diffus wachsende Formen ohne Mikrokalk
Lobuläres Karzinom	Häufig späte Entdeckung aufgrund diffusen Wachstums, fehlendem Mikrokalk, selten Opazitäten

1.3 Verkalkungen

1.3.1 Verkalkungen in der Mammographie

Wie Eingangs beschrieben liegen die Stärken der Mammographie in der hohen Sensitivität für Mikrokalk ab einem Durchmesser von 10 mm [60].

Wenn auch nicht spezifisch für maligne Veränderungen so sind Verkalkungen, speziell die Mikroverkalkungen, doch häufig damit in Verbindung zu bringen.

Die in-situ- Karzinome, insbesondere das DCIS, weisen in bis zu 85% der Fälle Kalk auf [65]. Bei den invasiven Karzinomen sind es hingegen nur 30% [66]. Da die frühen Tumorstadien meist noch nicht einer Palpation zugänglich sind, kommt hier der Mammographie eine die entscheidende Rolle in der Früherkennung zu [67, 68].

Die Verkalkungen spiegeln in der Regel das Produkt eines sich in der Brustdrüse entwickelnden Prozesses wider. Die vorrangigen Ziele bei der Beurteilung der Verkalkungen sind die frühe Erfassung dieser Prozesse und die Lokalisation der genauen anatomischen Entstehungsstelle.

Als pathogenetische Ursachen kommen verschiedene Möglichkeiten in Betracht. Die Verkalkungen können einerseits aus nekrotischem Zellmaterial, infolge degenerativer Prozesse, Entzündungen und Narbenbildung bedingt sein. Andererseits bewirken auch Parasiten und Fremdkörper sowie durch

Kristallisation eingedicktes, zelluläres Sekret die Entstehung von Kalkeinlagerungen.

Bei der Beurteilung des Kalks muss die Anordnung bzw. die Verteilung beachtet werden. Verkalkungen können diffus über die ganze Brust verteilt sein, sie können in Grüppchen eng beieinander liegen oder segmental und lobulär lokalisiert sein. Abhängig von Form, Größe und Abstand untereinander kann so die Dignität abgeschätzt werden.

Genau hierin liegt jedoch die Schwierigkeit für die beurteilenden Radiologen. Obwohl sich viele Studien mit der radiologisch-histologischen Korrelation beschäftigt haben, bleibt immer noch eine große Ungewissheit zwischen der Beziehung von radiologischen und pathologischen Charakteristiken der Verkalkungen [69].

Im Folgenden werden die anerkannten Charakteristika von Mikro- und Makroverkalkungen aufgeführt.

1.3.2 Makroverkalkungen

Ein wichtiger Faktor bei der Beurteilung der Dignität von Verkalkungen ist die Größe. Verkalkungen über 2 mm im Durchmesser werden als Makroverkalkungen bezeichnet. Diese sind in der Regel gutartig. Solche benignen Makroverkalkungen können sekretorisch bei der Plasmazellmastitis, bei der Liponekrose, bei Hautverkalkungen, Gefäßverkalkungen und Talgdrüsenverkalkungen vorkommen. Außerdem findet man Makroverkalkungen beim Fibroadenom, Papillom, Hämangiom und bei Narben- und Lymphknotenverkalkungen.

1.3.3 Mikroverkalkungen

Als Mikroverkalkungen werden Verkalkungen bezeichnet, die kleiner als 1 mm sind. Mikroverkalkungen können auf einen malignen Prozess hinweisen, aber

ebenfalls Ausdruck einer benignen Veränderung sein.

Benigne Mikroverkalkungen sind lobuläre Verkalkungen wie Sekret- und Involutionsverkalkungen. Die im Alter auftretenden Involutionsverkalkungen entstehen im atrophischen Lobulus und kommen multipel vor. Die lobulären Verkalkungen kommen bei der sklerosierenden Adenose und bei erhöhtem Calciumgehalt als so genannte Kalkmilch-Zysten vor.

Maligne Mikroverkalkungen sind fast immer Ausdruck eines Duktalen Karzinoms. Wenn zusätzlich Verdichtungen zu sehen sind, kann dies auf ein invasives Wachstum hinweisen.

Ein weiteres wichtiges Beurteilungskriterium neben der Größe ist die Form der Verkalkungen. So zeigen sich die typisch benignen Verkalkungen wie z.B. die Involutionsverkalkungen als puderartige, runde Verkalkungen. Bei der Kalkmilch sieht man bei vertikalem Strahlengang runde Verkalkungen. Bei horizontalem Strahlengang sind halbrunde, sichelförmige Verkalkungen erkennbar. Beim Fibroadenom sind die Verkalkungen grobschollig und polymorph. Bei der Plasmazellmastitis stellen sich die Verkalkungen linear, gleichmäßig dicht, mit glatter Kontur und überwiegend nicht verzweigt dar[70].

In dieser Gruppe gibt es einerseits grobgranuläre Mikroverkalkungen, die im nekrotischen Zentrum eines malignen Prozesses im Gangsystem entstehen und andererseits feingranuläre Mikroverkalkungen. Diese entstehen in den Lobuli durch eingedicktes Sekret zwischen den Tumorzellen und durch Kristallisation. Bei letzteren gestaltet sich die Abgrenzung zu benignen Veränderungen besonders schwierig.

Die grobgranulären, duktaalen Verkalkungen haben meist eine linear-verzweigende Anordnung und sind Ausdruck eines wenig oder mäßig differenzierten DCIS. Das Risiko einer Invasion ist hier groß.

Die feingranulären, lobulären Verkalkungen sind meist perlschnurartig angeordnet und deuten auf ein gut differenziertes DCIS hin. Das Risiko einer Invasion ist hierbei klein.

Zusammenfassend lässt sich sagen, dass eine genaue Analyse solcher Verkalkungen mit differenzialdiagnostischen Überlegungen unverzichtbar und notwendig ist für das weitere therapeutische Vorgehen.

Tab. 2: Gegenüberstellung ‚maligner - ‚benigner‘ Verkalkungen in der Mammographie

	Eher typisch Benigne	Eher typisch Maligne
Form	Rund, oval, glatte Ränder, Wände nicht maskiert, scharf begrenzt, grobgranulär, ‚pop-cornartig‘, rutenförmig, zentrale Aufhellung mit stärker verkalktem Rand, ringförmig, ‚Teetassenphänomen‘.	Fein linear, duktal, fein verästelnd, y-förmig, pleomorph, irregulär, feingranulär, heterogen-granulär, amorph, unscharf begrenzter Rand.
Größe	Größer (Makroverkalkungen)	Kleiner (Mikroverkalkungen).
Verteilung	Diffus (regional, gruppiert) selten auch segmental)	Linear (segmental, regional) gruppiert)

1.4 BI-RADS

1.4.1 Allgemeines

Das ‚Breast Imaging Reporting and Data System‘, kurz BI-RADS, des ‚American College of Radiology‘ (ACR) wurde 1986 entwickelt. Ziel war es, die Qualität der Beurteilung von Mammographien sicherzustellen und somit eine verbesserte Patientenbehandlung zu gewährleisten.

Umfangreiche Qualitätssicherungen in den USA führte in den Achtziger Jahren zur Entwicklung eines Mammographie-Akkreditierungs-Programms des ACR [71, 72]. Dieses Programm fördert die Zertifizierung der mammographischen Ausrüstung, die Schulungsanforderungen für Radiologen und technischem Personal sowie die Standards für Qualitätskontrollen.

Ein besonderes Problem war bis dahin die unterschiedliche Terminologie innerhalb der mammographischen Befundung [48, 73-76]. Daher entwickelte

das Lexicon Committee des ACR das BI-RADS® System. Durch diese Standardisierung der Mammadiagnostik mit Hilfe von BI-RADS®, werden interindividuelle Unterschiede der subjektiven Bildinterpretation verringert.

Die vierte, aktuelle Ausgabe des BI-RADS®- Atlanten stammt aus dem Jahre 2003 [70].

Im ersten der insgesamt sechs Kapitel des BI-RADS- Atlanten, dem „Breast Imaging Lexicon“, legt das ACR Begriffe und Definitionen fest. Damit kategorisiert es alle krankhaften und nicht krankhaften Brustveränderungen mit dem Ziel einer klaren, prägnanten und standardisierten Terminologie.

Die mammographischen Veränderungen werden in die Untergruppen Opazität, Verkalkung, Architekturstörung und Spezialfälle eingeteilt. Die Befunde werden nach Größe, Form und Begrenzung beschrieben. Außerdem werden bei der Beschreibung der Brustveränderungen auch deren Lokalisation und assoziierte Befunde wie Hautverdickungen, Haut- und Brustwarzenretraktion oder eine axilläre Adenopathie mit einbezogen.

Schließlich muss bei der endgültigen Bewertung der mammographisch erfassten Läsion einkalkuliert werden, dass die diagnostische Genauigkeit invers mit der Parenchymdichte der Brustdrüse korreliert. Je dichter, desto schlechter beurteilbar und desto leichter kann ein Befund übersehen werden. Das ACR teilt die Dichte des Brustdrüsengewebes daher in 4 Kategorien ein, von einem lipomatösen und gut beurteilbaren Typ I bis zu einem sehr dichten und begrenzt beurteilbaren Typ IV.

Nach BI-RADS werden die erhobenen Befunde wie in Tabelle 3 ersichtlich in sieben Kategorien eingeteilt. Durch diese Einordnung werden die Wahrscheinlichkeit für das Vorliegen einer malignen Läsion und das weitere diagnostische und therapeutische Vorgehen festgelegt.

Obwohl es keine präbiopische Untersuchung zum sicheren Ausschluss von Brustkrebs gibt, unterstützt und hilft BI-RADS bei der Einordnung und Risikoabschätzung der mammographischen Befunde beziehungsweise der

Erkrankungswahrscheinlichkeit [73].

Tab. 3: Übersicht BI-RADS- System [70, 77]

Kategorie	BIRADS- Bewertung	Procedere	Karzinomwahrscheinlichkeit
0	Diagnostik nicht komplett	Komplettierung	Keine Aussage
1	Kein pathologischer Befund	Keine Kontrolle	<1%
2	Sicher benigner Befund	Keine Kontrolle	<1%
3	Wahrscheinlich benigner Befund	Kontrolle in 6 Monaten	1-3%
4 4a 4b 4c	Unklarer, wahrscheinlich maligner Befund	PE/Biopsie	20-90% 4-20% 20-50% 50-90%
5	Malignom- typischer Befund	Biopsie + adäquate multimediale Therapie	>90%
6	Maligner, histologisch nachgewiesener Befund	Adäquate multimediale Therapie	100%

1.4.2 Was sagt BI-RADS über Verkalkungen?

Gutartige Verkalkungen sind in der Regel größer als solche Verkalkungen die mit Malignität assoziiert sind. Sie sind meist größer, oft rund mit gleichmäßigeren Rändern und leichter erkennbar. Mit Malignität assoziierte Verkalkungen sind in der Regel klein, flau und oft erst durch Vergrößerung

erkennbar.

Das BI-RADS– Lexikon [70, 77] unterscheidet drei verschiedene Typen von Verkalkungen: Typisch Gutartige, mittelgradig Suspekte und Verkalkungen mit einer höheren Wahrscheinlichkeit von Malignität. Im Einzelnen werden darunter verstanden:

1. TYPISCH GUTARTIG

a) Hautverkalkungen:

Pathognomische Ablagerungen mit einem typisch transparentem Zentrum.

b) Vaskuläre Verkalkungen:

Parallele Bahnen oder linear tubuläre Verkalkungen, die Blutgefäßen zuordbar sind.

c) Grobe, korkenzieherartige oder „popcorn-ähnliche“ Verkalkungen:

Klassische Verkalkungen, die durch ein in Involution befindliches Fibroadenom hervorgerufen werden.

d) Große astartige Verkalkungen:

Verkalkungen, die zusammenhängende Äste bilden und in der Regel einen Durchmesser von mehr als 1 mm besitzen. Diese Art von Verkalkungen findet sich bei sekretorischen Erkrankungen, Plasmazellmastitis und Duktektasien.

e) Rundliche Verkalkungen:

Häufig multipel auftretende und größenvariable Verkalkungen, die sich in den Azini der Lobuli bilden.

f) Verkalkungen mit transparentem Zentrum:

Diese Ablagerungen haben eine glatte Oberfläche, sind rund oder oval und weisen ein transparentes Zentrum auf.

g) Verkalkungen vom „Eierschalen“-Typ:

Sehr dünne Verkalkungen, die als Kalkablagerungen auf der Oberfläche einer Kugel erscheinen.

h) Kalkmilch-Verkalkungen oder Teetassen-Verkalkungen:

Diese werden durch Kalksedimente in Zysten hervorgerufen.

i) Nahtverkalkungen:

Hier wird auf Nahtmaterial abgelagerter Kalk gebildet. Sie imponieren als typisch lineare oder tubuläre Verkalkungen.

j) Dystrophische Verkalkungen:

Diese Verkalkungen bilden sich häufig in der bestrahlten Brust oder im Anschluss an ein Brusttrauma. Häufig findet man ein transparentes Zentrum.

k) Punktierte Verkalkungen:

Diese sind rund oder oval, kleiner als 0,5 mm und weisen gut definierte Ränder auf.

2. MITTELGRADIG SUSPEKTE VERKALKUNGEN

a) Amorphe oder unscharfe Verkalkungen:

Runde oder flockige Verkalkungen, die in ihrem Erscheinungsbild derart klein oder verschwommen sind, dass eine genauere morphologische Klassifizierung nicht vorgenommen werden kann.

3. HÖHERE WAHRSCHEINLICHKEIT VON MALIGNITÄT

a) Pleomorphe oder heterogene, granuläre Verkalkungen:

Diese sind gewöhnlich irreguläre Verkalkungen variabler Größe und Form und kleiner als 0,5 mm im Durchmesser. Sie sind suspekter als

amorphe Formen und weder typisch gutartig noch typisch bösartig.

b) Feine, lineare, eventuell verästelte Verkalkungen:

Dabei handelt es sich um dünne, irreguläre Verkalkungen, die linear erscheinen, aber diskontinuierlich sind und im Durchmesser kleiner als 0,5 mm sind.

Außerdem wird im BI-RADS[®] Lexikon das Verteilungsmuster beschrieben. Dabei wird folgende Einteilung vorgenommen:

a) Gruppiert:

Multiple Verkalkungen in einem kleinen Gewebsvolumen. Je dichter und je mehr Verkalkungen desto wahrscheinlicher wird ein Karzinom

b) Linear:

Verkalkungen, die in einer Linie angeordnet sind. Aufzweigungen sind möglich. Lineare Verkalkungen weisen auf Ablagerungen in Milchgängen hin und können Ausdruck eines intraduktalen malignen Prozesses sein.

c) Segmental:

Auch diese Verkalkungen sind insofern suspekt, als ihr Verteilungsmuster auf Ablagerungen in Milchgängen und deren Verzweigungen hinweist. Es erhöht sich die Wahrscheinlichkeit für das Auftreten eines multifokalen Karzinoms in einem Lobus oder einem Segment der Brust. Obwohl es gutartige Ursachen für segmentale Verkalkungen, wie beispielsweise sekretorische Erkrankungen gibt, ist dieses Verteilungsmuster umso verdächtiger, wenn die Morphologie der Verkalkungen nicht typisch gutartig ist.

d) Regional:

Diese eher gutartigen Verkalkungen sind in einem großen Teil des Brustgewebes verstreut. Dabei ordnen sie sich nicht unbedingt in einem duktalem Verteilungsmuster an. Sie finden sich aber nicht überall in der Brust

und sind den übrigen, eher suspekten Kategorien nicht zuzuordnen.

e) Diffus / Verstreut:

Diese Verkalkungen sind zufällig in der gesamten Brust verteilt und deuten eher auf eine gutartige Diagnose hin.

Die Einteilung von Mikroverkalkungen und ihre Zuordnung zu den BI-RADS- Klassen bereiten trotz den oben genannten Definitionsversuchen des ACR Schwierigkeiten. Auch findet man nur wenige Studien über die Klassifikation von Mikroverkalkungen nach dem BI-RADS Lexikon [78-80].

Im Katalog des ACR wird zwar erläutert, welche mammographischen Veränderungen eher für Benignität oder für Malignität sprechen, es wird jedoch nicht explizit hervorgehoben in welche Kategorie der Befund dann eingeordnet werden soll [81].

Die Kategorieeinordnung muss der Befunder anhand aller erfassten mammographischen Veränderungen ganzheitlich, auch im dynamischen Zusammenhang gesehen, abschätzen. Die Detektion eines malignen Tumors hängt von einer Vielzahl von Faktoren ab. Dazu zählen die Größe der Läsion, deren Auffälligkeit, die Dichte des Brustdrüsengewebes, das Alter der Patientin und ihr Hormonstatus, die Bildqualität und nicht zuletzt der Interpretativen Fähigkeiten und die Erfahrung des Radiologen [82, 83].

Dass die subjektive Erfahrung eine große Rolle spielt, zeigen viele Studien, die belegen, dass trotz der Standardisierung durch BI-RADS immer noch eine hohe Variabilität zwischen verschiedenen Befundern vorhanden ist [28, 84].

Viele Befunder kommen also trotz Standardisierung durch das BI-RADS Lexikon zu verschiedenen Ergebnissen und somit zu verschiedenen BI-RADS Einteilungen was Auswirkungen auf die Prognose für die betroffenen Patienten hat.

In der aktuellen vierten Ausgabe des BI-RADS Lexikon [70] ist zwar die

kombinierte Beurteilung von Morphologie und Verteilung von Mikroverkalkungen vorgeschlagen es werden aber keine eindeutige Regeln für eine solche Einschätzung definiert [81].

Weil diese Zuordnung von Mikrokalk zu den Kategorien selbst im Katalog des ACR nicht eindeutig definiert ist, wird an der Universitäts-Frauenklinik in Tübingen ein, vor dem Jahre 2000 dort entwickeltes Mikrokalkmatrixsystem, verwendet. Bisher ist aber noch nichts darüber publiziert worden.

Die anhand der Matrix vorgeschlagene Aufarbeitung von Mikroverkalkungen anhand von Morphologie und Verteilung ist ein Versuch die ‚Lücke‘ zwischen der Beschreibung von Verkalkungen und die Einteilung zu einer BI-RADS Kategorie zu überbrücken. [81]

1.5 Aufgabenstellung und Zielsetzung

Ziel dieser Arbeit ist die Untersuchung der Möglichkeit einer Verbesserung der Klassifikation von Mikroverkalkungen in Mammographien mit den BI-RADS® Kategorien durch ein Matrix-System.

Die zentrale Frage stellt sich nach der Möglichkeit, Mikroverkalkungen mit einem neuen Matrix-System so zu klassifizieren, dass auch ein noch unerfahrener Befunder Zusammenhänge mit den BI-RADS- Kategorien erkennen kann und dass die Zuordnung zu einer Kategorie leichter fällt.

Für die Beurteilung der Klassifizierungen wird die histologische Aufarbeitung der fraglichen Areale hinzugezogen.

Folgende Aspekte werden diskutiert:

- Ist die Benutzung des Matrixsystems sinnvoll? Wie ist seine Sensitivität bzw. Spezifität im Vergleich zur subjektiven Beurteilung? Worin liegt der eventuell gefundene Nutzen?
- Kann die Einteilung zu einer BI-RADS- Kategorie mit Hilfe der Matrix bei

vorhandenem Mikrokalk erleichtert werden?

- Ist die Zuordnung zu den BI-RADS Klassen in der Matrix sinnvoll? Entsprechen die zugeteilten BI-RADS Klassen auch im Hinblick auf die Malignitätswahrscheinlichkeit den tatsächlichen Gegebenheiten?
- Kann das Schema vor allem für noch unerfahrene Befunder eine Hilfestellung bieten?

Zur Erfassung der Daten dient eine retrospektive ‚multi-reader‘ Studie. Die insgesamt fünf Untersucher dieser Studie haben unterschiedlich viel Erfahrung im Bereich der Befundung von Mammographien. Sie werteten Mikroverkalkungen in Mammographien von 203 Patientinnen nach den BI-RADS Kategorien aus. Zum einen mit Hilfe des Matrix-Schemas und zum anderen nach ihrer subjektiven Einschätzung.

2. Material und Methoden

2.1 Patienten

In die Studie wurden zu Beginn 280 Patientinnen aufgenommen. Bei allen Frauen wurde im Zeitraum vom Jahr 2002 bis zum Jahr 2004 im Brustzentrum der Universitätsfrauenklinik Tübingen und der Radiologischen Universitätsklinik eine mammographische Diagnostik und Biopsie der Brust durchgeführt.

Die Entscheidung für die Aufnahme in die Studie wurde anhand des Pathologischen Befundes getroffen. Laut Einschlusskriterien musste in diesem histologisch Kalk nachgewiesen sein.

Es handelt sich um ein heterogenes Patientenkollektiv, da die individuelle Indikation zur Mammographie aus verschiedenen Gründen gestellt wurde:

- Screening bei unauffälligem Untersuchungsbefund.
- Verdächtiger Tastbefund oder Brustläsion.
- ‚Follow-up‘ Untersuchung bei Zustand nach Karzinom oder ‚Carcinoma in situ‘.
- Screening bei kontralateralem Karzinom.
- ‚Follow-up‘ Untersuchung bei BI-RADS 3 Läsion.

Das Alter der untersuchten Patientinnen lag zwischen 22 und 79 Jahren. Dabei lag das Durchschnittsalter bei etwa 56,5 Jahren.

Für die Aufnahme in die Studie bzw. den Verbleib in der Studie galten folgende Kriterien:

Einschlusskriterien:

1. Mikrokalk in der mammographischen Untersuchung.
2. Entnahme einer Biopsie des Mikrokalkareals mit histologischer Aufarbeitung.
3. Gefundener Mikrokalk in der Histologie.

4. Der histologisch gefundene Mikrokalk ist eindeutig mit dem Mikrokalkareal der Mammographie in Verbindung zu bringen.
5. Digitale Mammographiebilder.

Ausschlusskriterien:

1. Eindeutige BI-RADS 5 Opazität in der Umgebung des Mikrokalkes.
2. Konventionelle Bilder (z.B. bei Fremdbefundung).
3. In der Histologie wurde kein Mikrokalk gefunden.
4. Der im Histologiebefund beschriebene Mikrokalk lässt sich nicht eindeutig mit dem befundeten Mikrokalk in der Mammographie in Verbindung bringen.
5. Keine ausreichende Vergrößerungsaufnahmen bzw. schlechte Beurteilungsbedingungen (z.B. bei ausgeprägter Architekturstörung oder zu dichtem Brustdrüsengewebe).
6. Biopsieentnahme aufgrund eines sonographisch suspekten Befundes bzw. suspekten MR- Befundes und nicht wegen gefundenen Mikrokalks in der Mammographie.

Im Verlauf der Studie mussten 77 Patienten wegen Nichteinhaltens einer oder mehrerer der oben genannten Kriterien ausgeschlossen werden. Zur statistischen Auswertung kamen daher noch 203 Patienten.

2.2 Diagnostische Verfahren

2.2.1 Mammographie

Die mammographische Untersuchung erfolgte in der Frauenklinik der Universität Tübingen. Die Mammographie wurde mit einem G.E. Senographe 2000 D (General Electric Medical Systems, Milwaukee, WI, USA) mit dazugehöriger Workstation erstellt. Es handelt sich um einen digitalen Vollfeld-Detektor mit amorphen Siliziumkristallen, einem Filmformat von 19x23 cm und

Belichtungsautomatik (25-32 Kilovolt (kV)). Der Quantenwirkungsgrad (= detective quantum efficiency, DQE) und die Kontrastauflösung sind hoch und erlauben eine für das menschlich Auge optimale räumliche Auflösung [59].

Außerdem ist eine Kompressions-Vergrößerungs-Einrichtung vorhanden, die bei nicht zweifelsfrei zu beurteilenden Befunden, wie z.B. überlagerte Mikroverkalkungen, diese Areale der Brust zur besseren Darstellung vergrößert. Zur Auswertung kamen nur digitale Bilder dieses Gerätes.

Die Mammographien erfolgten jeweils in den Standardebenen (kraniokaudaler und mediolateral-obliquier Strahlengang). Des Weiteren wurde bei einigen Patientinnen bei unklaren Befunden Zusatzaufnahmen wie beispielsweise eine Schrägprojektion (mediolateral) durchgeführt.

Die Qualität der technischen Aufnahmen orientierte sich an den europäischen Leitlinien [85].

2.2.2 Biopsieverfahren und Histologie

Bei allen Mammographien der in die Studie aufgenommenen Patientinnen wurden Mikrokalkareale bzw. Mikroverkalkungen gefunden, die einer Biopsie zugeführt wurden.

Es kamen zwei verschiedene Biopsiearten zur Anwendung, jeweils entweder sonographisch oder stereotaktisch digital gesteuert:

- Offene Biopsie bei brusterhaltender Therapie (BET) oder Ablatio
- Vakuumbiopsie (VB)

Die Vakuumbiopsien wurden alle am Mammothestsystem der Firma Fischer (= Fischertisch) in der Radiologie der Frauenklinik gemacht. Der Fischertisch ist ein Stereotaxietisch mit integrierter digitaler Mammographieröhre und Biopsieeinheit. Es wurden jeweils zwischen 12-36 Gewebeproben aus dem suspektem Areal entnommen.

Die offenen Biopsien wurden nach mammographischer Drahtmarkierung bzw.

nach ‚coil‘ -Markierung bei BET durchgeführt.

Feinnadelaspirationen kamen nicht zur Anwendung, da diese bei Mikroverkalkungen nicht indiziert sind.

Von den Gewebeproben, die durch VB gewonnen wurden, wurde zur Überprüfung der erfolgreichen Punktion eine Röntgenaufnahme, eine so genannte Präparateradiographie gemacht. Mit der Präparateradiographie kann sichergestellt werden, dass das biopsierte Areal auch tatsächlich den verdächtigen Mikrokalk enthält.

Die histologische Aufarbeitung und Diagnostik der Präparate wurden im Pathologischen Institut der Universität Tübingen unter Leitung von Herrn Prof. Dr. B. Bültmann angefertigt.

2.3 Aufbereitung und Auswertung der Patientendaten

2.3.1 Aufbereitung der Daten

Folgende Daten der 280 untersuchten Patientinnen wurden erfasst und mittels des Datenbankprogramms Access© (Microsoft) katalogisiert:

1. Patientename
2. Geburtsdatum
3. Brustseite
4. Patientennummer
5. ‚IOP‘-Nummer (=Anonymisierungsnummer)
6. Mammographie-Untersuchungsdatum
7. Histologischer Befund der Biopsie oder der Resektion bei OP
8. Datum der Biopsieentnahme bzw. der OP
9. Zweitkarzinom (ja/nein)
10. Digitale Bilder (ja/nein)
11. Maligne Histologie (ja/nein)

2.3.2 Befundung der Mammographien

Die digitalen Mammographien der in die Studie aufgenommenen Patientinnen wurden automatisch an der Befund- Workstation anonymisiert und von vier Radiologen an der Radiologischen Universitätsklinik beschäftigten Radiologen unabhängig voneinander zweimal befundet. Diese hatten unterschiedlich viele Jahre Erfahrung in der Befundung von Mammographien:

Reader 1: 5 Jahre

Reader 2: 7 Jahre

Reader 3: 6 Jahre

Reader 4: 4 Jahre

Um einen Erinnerungseffekt zu verringern wurde die zweite Befundung in einem Mindestabstand von 4 Wochen zum ersten Durchgang durchgeführt.

Waren in einer Mammographie mehrere unterschiedliche Mikrokalkareale vorhanden, wurde vereinbart immer das suspektteste Areal zu befunden. Bei unklaren Fällen wurde das zu untersuchende Areal vorher digital markiert um sicherzustellen, dass das korrekte und nachweislich biopsierte Gebiet befundet wird. So konnte bei diesen unklaren Fällen auch sichergestellt werden, dass bei der erneuten Befundung wieder dieselbe Gruppe Mikrokalk begutachtet wurde.

Ausnahme: Ein nicht in Tübingen beschäftigter Radiologe, mit der längsten Erfahrung von 12 Jahren, befundete 27 Mammographien jeweils einmal. Der Vollständigkeit halber geht er als Untersucher mit der Nummer 5 in die Auswertung mit ein, wird im Text jedoch nicht weiter aufgeführt.

Alle Mammographien wurden an der Review-Workstation des G.E. Senographe 2000D ausgewertet. Folgende Funktionen sind an der Workstation möglich:

- Vergrößerung
- Bildumkehr

- Messungen
- Veränderung der Helligkeits- und Kontrastunterschiede durch Fensterung

Die Einteilung in die Gruppenform bzw. die Einteilung des Einzelkalkes erfolgte in Anlehnung an die Erläuterungen im BI-RADS Lexikon und gemäß Tab.2 im Kapitel 1.3.3. Die nachgewiesenen Mikroverkalkungen wurden folgendermaßen charakterisiert. Dabei stellen die Zahlen in den Klammern die Umcodierung für die Statistische Auswertung dar:

1. Gruppenform:

Diffus (1):

Gleichmäßige Verteilung der Verkalkung über das gesamte Brustparenchym

Rund/Oval (2):

Gruppiertes Mikrokalk, in der Regel mit einem Volumen von 1ccm.

Regional (3):

Gruppiertes Mikrokalk, in der Regel mit einem Volumen von über 1 ccm.

Segmental (4):

Verkalkungen in einem zum duktalem System gehörenden Segment

Linear (5):

Duktale Verteilung, auch mit Verzweigungen. Typischerweise zur Brustwarze hinorientierender Kalk.

Die Verteilungsmuster wurden übereinstimmend mit dem BI-RADS Lexikon erarbeitet [70, 77]. Im Lexikon wird der Einfluss der unterschiedlichen Verteilungen auf die Einteilungskategorien weder definiert noch quantifiziert [81]. Die Zuordnung zu den unterschiedlichen Malignitätswahrscheinlichkeiten basiert daher auf Daten der Literatur [86-91].

2. Einzelkalk:

Typisch benigne:

Rund, abgerundet, Stäbchen, radiolucent, popcornartig, etc. (1)

Intermediär:

granulär etc. (2)

Typisch maligne:

polymorph, pleomorph, etc. (3)

Die Einteilung in einer der drei Gruppen erfolgte gemäß [70, 77].

3. Subjektive BI-RADS Einteilung: 0, 1, 2, 3, 4, 5 (0-5)

4. Opazität vorhanden?: ja/nein

5. Wenn ja, BI-RADS Kategorie der Opazität: 0, 1, 2, 3, 4, 5 (0-5)

2.3.3 Das Mikrokalkmatrixsystem

Das Mikrokalkmatrixsystem ist eine Interpretationshilfe für Mikroverkalkungen [81].

Sie enthält, wie in Kapitel 2.3.2 erläutert, klar definierte Verkalkungsmorphologien und klar definierte Verteilungsmuster der Verkalkungen.

Die Zuordnung zu den BI-RADS Kategorien zeigt Tabelle 4.

Tab. 4: Matrixsystem (* **Modifikation des Matrixsystems**)

	Einzelkalk		
Gruppe	Typisch benigne: Rund, Kurvilinear, Glattlinear	Intermediär: Granulär, Amorph	Typisch maligne: Verzweigend, Polymorph
Diffus	2	3	4
Rund/Oval	2	3 / 4 *	4
Regional	2	3	4
Segmental	3	4	5
Linear	3	5	5

2.3.4 Auswertung

Die Mammographien wurden von den vier Radiologen der Universität Tübingen (Ausnahme siehe oben) je zweimal anhand des Auswertebogens befundet.

Dadurch kam es zu einer Datensatzmenge (mittels einer Excel© (Microsoft) Tabelle katalogisiert) von 1665 Daten mit folgenden Informationen:

- 1.) Befundername bzw. Befundnummer.
- 2.) 1. oder 2. Reading.

3.) Einteilung nach der Matrix (2 Zahlenwerte: von 1. der Gruppe (1-5) und 2. des Einzelkalkes (1-3)) mit daraus resultierender BI-RADS- Kategorie (Kategorie: 2-5).

4.) Einteilung nach der modifizierten Matrix mit daraus resultierender BI-RADS- Kategorie.

5.) BI-RADS –Kategorie nach der Subjektiven Einteilung .

Bei der Einteilung mit der Matrix durften im Gegensatz zur Subjektiven Einteilung keinerlei mammographische Nebenbefunde wie Gewebsverdichtungen oder Hauteinziehungen beachtet werden. Die Matrix teilt der Mammographie allein aufgrund von Morphologie und Verteilung der Verkalkungen eine Kategorie zu.

Das Vermerken einer Opazität und deren BI-RADS- Kategorie bzw. das Vorhandenseins eines Zweitkarzinoms spielte weniger für die Auswertung eine Rolle als für das Überprüfen, ob die Mammographie für den Verbleib in der Studie geeignet ist.

Die Histologischen Befunde wurden entsprechend den Ein- und Ausschlusskriterien analysiert und die genaue histologische Diagnose vermerkt.

2.4 Statistische Erhebungen

Die statistische Auswertung der 203 in der Studie verbliebenen Patientinnen fand in Zusammenarbeit mit dem Institut für Medizinische Biometrie der Universität Tübingen statt (Direktor: Herr Prof. Dr. K. Dietz, Auswertung durch Frau Dr. Tina Herberts).

Als abhängige Variablen galten:

1.) Auswertemethode:

1. Auswertemethode: Matrixschema
2. Auswertemethode: Subjektive Beurteilung
3. Auswertemethode: modifiziertes Matrixschema

2.) Reader bzw. dessen Jahre an Erfahrung

3.) Reading (1. bzw. 2. Auswertung)

4.) Malignität (wahr/falsch): BI-RADS Kategorie 0-3 wurde als benigne, 4-6 als maligne eingestuft.

Von allen drei Auswertemethoden wurden jeweils die Spezifität und die Sensitivität berechnet:

Die Spezifität ist ein Maß für die Eindeutigkeit eines Testes. Sie ist die Wahrscheinlichkeit eines negativen Resultates (-T) bei gleichzeitigem Fehlen der Krankheit (K-) und wird bezeichnet mit $P(T-/K-)$. Sie stellt also den Anteil der zutreffenden als negativ diagnostizierten unter allen tatsächlich Gesunden dar.

Analog ist die Sensitivität ein Maß für die Empfindlichkeit eines Testes. Diese stellt den Anteil der zutreffenden als positiv diagnostizierten unter allen tatsächlichen Kranken dar, und wird bezeichnet mit $P(T+/K+)$

Um Sensitivität und Spezifität bestimmen zu können, wurde folgende Bezeichnung vorgenommen:

Alle unauffälligen, benignen Befunde wurden als „negativ“ (=0) bewertet. Alle malignen Befunde wurden als „positiv“ (=1) bewertet.

Das 1. und 2. Reading wurde, jeweils mit allen drei Methoden, bezüglich der Inter- und Intrareadervariabilität verglichen.

Außerdem wurde die Matrixeinteilung mit der subjektiven Einteilung je Untersucher und der Einteilung nach der modifizierten Matrix verglichen und die Übereinstimmung der malignen Befunde insgesamt erhoben.

Die Gesamteinteilungen in eine BI-RADS Kategorie und Karzinomanteile je Kategorie wurden für die unterschiedlichen Methoden und Reader errechnet.

Als letzte statistische Auswertung erfolgte die Berechnung der Malignitätswahrscheinlichkeit über die Matrix.

Als statistische Tests kamen der Chi-Quadrat-Test, der Likelihood-Ratio-Test, der Pearson-Test, und der Wald-Test zur Verwendung. Als signifikant wurde ein Ergebnis bei einem p-Wert $< 0,05$ angesehen. Die Modellgüte wurde mit dem Whole-Modell-Test berechnet.

3. Ergebnisse

3.1. Histologisch-Pathologische Ergebnisse

Da es sich um eine retrospektive Studie handelt, lagen bei Beginn der Datenerhebung bereits alle histologischen Ergebnisse vor.

Aus den pathologischen Befunden wurden jeweils die wichtigen Hauptdiagnosen aufgenommen. Diese sind in Tabelle 5 nach der Häufigkeit aufgelistet.

Tab. 5: Histologische Diagnosen und Fallanzahlen

DIAGNOSE	ANZAHL	DIAGNOSE	ANZAHL
DCIS	89	LIN 1 (ALH)	2
Mastopathie (gesamt)	58	Fibroadenomatöse Hyperplasie	1
Davon: Fibrös-zystisch	49		
Fibrös	9		
Adenosen (gesamt)	42	Tubuloductales Karzinom	1
Davon: Sklerosierend	33		
Tubulär	3		
IDC	35	Drüsengewebe und Gefäßkalk	1
Davon: DCIS mit IDC	22	Lymphangiosis carcinomatosa	1
Duktale Hyperplasie (DH)	25		
Apokrine Metaplasie	13	Im Fettgewebe auskristallisierter Mikrokalk	1
ILC	9	Grobschollig dystropher Mikrokalk	1
(Duktale) Papillome	7	Invasives muzinöses Adenokarzinom	1
Fibroadenom	6	Medulläres Karzinom	1
Atypisch ductale Hyperplasie (ADH)	4	Tubulopapilläres Adenom	1
Chronische Galaktophoretis	3	Atrophie	1
Ductulolobuläres Karzinom	2	Fibrose	1
Narbgewebe	2	Fettgewebsnekrosen	1
lobuläre Hyperplasie	2	Adenom	1
Intraductale Epithelproliferation	2	Fibrohyalinose	1

Ergebnisse

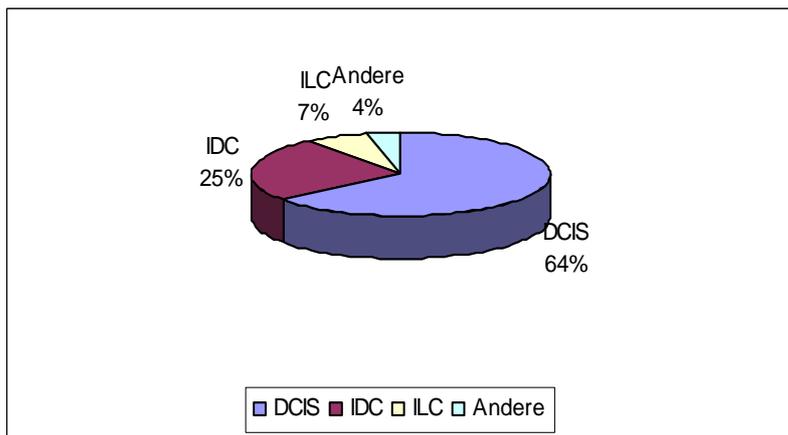


Abb. 3: Anteil maligner Diagnosen

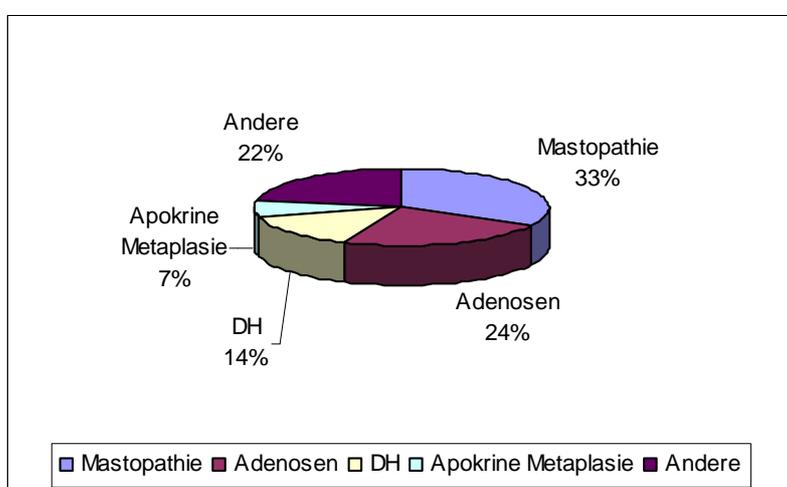


Abb. 4: Anteil benigner Diagnosen

Es wurden nur solche Diagnosen aufgenommen, die mammographisch eindeutig mit dem Areal des entnommenen Mikrokalkes in Verbindung stehen. Außerdem musste in der histologischen Aufarbeitung Mikrokalk gefunden werden. Dieser ging nicht gesondert in die Diagnoseliste mit ein, da sein Vorkommen ein Einschlusskriterium in die Studie war.

Insgesamt sind 312 Diagnosen in Tabelle 5 aufgeführt. Bei einer Ausgangszahl von 203 Patientinnen bedeutet dies, dass viele Patientinnen mit mehreren Diagnosen in Tabelle 5 vertreten sind.

Bei 90 Patientinnen sind zwei oder mehr Hauptdiagnosen aufgenommen. Genaue Zahlen sind in Tabelle 6 zu sehen. Bei 113 Patientinnen ist das

Mikrokalkareal mit nur einer Hauptdiagnose in Verbindung zu bringen. Zu einem großen Anteil ist dies eine Karzinomdiagnose. Insbesondere das DCIS, das IDC und das ILC.

Tab. 6: Gegenüberstellung Anzahl Diagnosen-Anzahl Patienten

ANZAHL DIAGNOSEN	ANZAHL PATIENTEN
1	113
2	72
3	17
4	1

Die benignen Veränderungen traten oft gemeinsam auf. Hier sind insbesondere die fibrös-zystische Mastopathie, die sklerosierende Adenose und die duktale Hyperplasie als die drei häufigsten benignen Diagnosen zu nennen. Abbildung 4 zeigt den Anteil der einzelnen benignen Diagnosen.

Das DCIS ist mit einer Anzahl von 89 Fällen sowohl die häufigste Diagnose wie auch die häufigste Karzinomdiagnose. Ein weiteres häufiges Karzinom ist das IDC. Das ILC kommt dagegen mit einer Anzahl von neun Fällen vergleichsweise selten vor. Das ductulolobuläre Karzinom, das Tubuloductale Karzinom, das invasive muzinöse Karzinom und das medulläre Karzinom sind Einzelfälle. Das LCIS ist nicht in der Diagnosenliste vertreten. Bei 2 Patientinnen wurde histologisch jedoch die Diagnose einer lobularen intraepithelialen Neoplasie Grad eins gestellt (=LIN 1) gestellt. Dies entspricht einer Atypisch lobulären Hyperplasie (ALH) und gilt als Vorstufe eines LCIS. Abbildung 3 gibt den Anteil der häufigsten malignen Diagnosen wieder.

3.2. Statistische Ergebnisse

3.2.1 Methode 1: Auswertung mit dem Matrixschema

3.2.1.1 Sensitivität

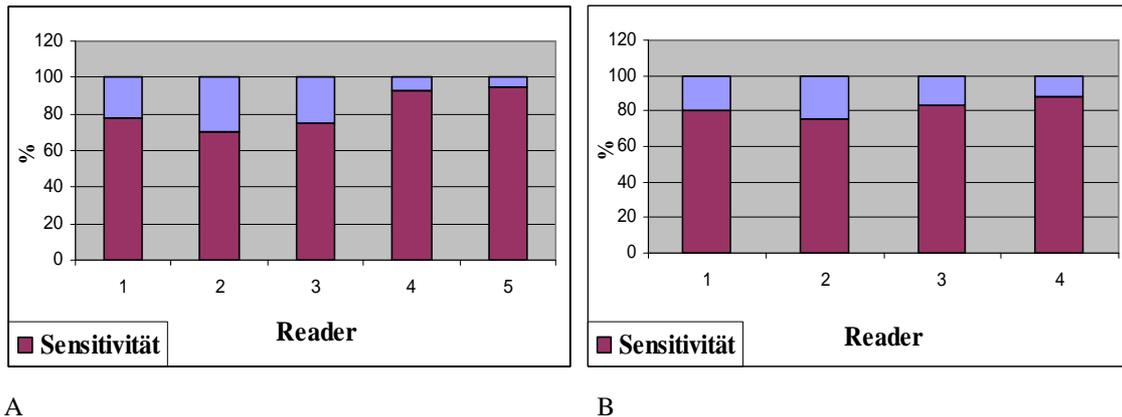


Abb. 5 :Sensitivität Methode 1. Angaben in Prozent, A: erstes Reading, B: zweites Reading.

Tab. 7: Sensitivität und falsch negative Befunde der Methode 1. Angaben in Prozent und Gesamtanzahl der Einteilungen (= Gesamt). 0= falsch negative Einteilungen, 1= richtig positive Einteilungen. Total= Gesamteinteilungen je Untersucher und je Dignität. A: erstes Reading, B: zweites Reading.

A				B			
Gesamt %	0	1	Total	Gesamt %	0 benigne	1 maligne	Total
Untersucher 1	22 21.57	80 78.43	102	Reader 1	20 19.61	82 80.39	102
Reader 2	31 30.39	71 69.61	102	Reader 2	25 24.51	77 75.49	102
Reader 3	26 25.49	76 74.51	102	Reader 3	17 16.67	85 83.33	102
Reader 4	7 6.86	95 93.14	102	Reader 4	12 11.76	90 88.24	102
Total	68	322	408	Total	74	334	408
Reader 5	1 5	19 95	20				
Total	87	341	428				

Wie man aus Tabelle 7 erkennen kann wurde bei 102, der insgesamt 203, ausgewerteten Patientinnen die Diagnose eines malignen Befundes gestellt.

Die Auswertung mit dem Matrixschema zeigte folgende Sensitivitäten:

Reader 1 hat in der ersten Auswertung 78,43% aller kranken Patienten als maligne eingestuft. Dies entspricht der Anzahl von 80 Patienten. Die restlichen 22 wurden von ihm als benigne kategorisiert. Im zweiten Durchgang beträgt die Sensitivität 80,39%. Hier wurden 82 Patienten richtig als krank erkannt und 20 für gesund eingeschätzt.

Untersucher 3 erreicht eine Sensitivität von 74,51% bei der ersten Befundung und 83,33% im Zweiten. Dies entspricht einer Gesamtanzahl von 76 bzw. 85 Patienten. Die Sensitivitäten von Untersucher 2 betragen 69,61% beziehungsweise 75,49%. Hier wurden 71 bzw. 77 Patienten mit malignen Befunden als krank kategorisiert.

Untersucher 4 erreicht beim der ersten Befundung eine Sensitivität von 93,14%. Er hat 95 der tatsächlich Kranken als krank, bzw. den Befund als maligne, erkannt. Sieben maligne Befunde wurden von ihm als benigne, der Patient als somit als gesund eingestuft. Bei der zweiten Befundung beträgt seine Sensitivität 88,24%. Dies entspricht einer Zahl von 90 Patienten, die richtig als krank kategorisiert wurden und 12 Patienten, die obwohl ein maligner Befund vorliegt, als gesund eingestuft wurden.

Abbildung 5 hebt die Unterschiede zwischen den Untersuchern nach Reading getrennt hervor.

Bei der ersten Auswertung wurde mit beiden Tests, wie in Tabelle 8 ersichtlich ist, eine Signifikanz der Sensitivität bewiesen. Dies wiederum zeigt, dass die Sensitivität zwischen den Readern signifikant unterschiedlich ist und damit der Untersucher Einfluss auf die Sensitivität hat. Die Ergebnisse für die zweite Auswertung sind dagegen nicht signifikant.

Ergebnisse

Tab. 8: Signifikanztests der Sensitivitätsberechnung bei der 1. Methode. A: erstes Reading, B: zweites Reading, W= Wahrscheinlichkeit, $W > \text{Chi}^2 =$ Signifikanzniveau.

A			B		
Test	Chi ²	W>Chi ²	Test	Chi ²	W>Chi ²
Likelihood Ratio	25.789	<.0001	Likelihood Ratio	5.977	0.1127
Pearson	22.475	0.0002	Pearson	5.877	0.1178

Aus Abbildung 6 und Tabelle 9 kann man die Gesamtsensitivität der Methode 1 erkennen. Sie beträgt 80,74%. Bei ihrer Berechnung werden alle Untersucher und beide Auswertungen zusammen betrachtet. Dadurch kommt es zu einer Datensatzmenge von insgesamt malignen 836 malignen Fällen, von denen 675 auch korrekt eingeteilt wurden.

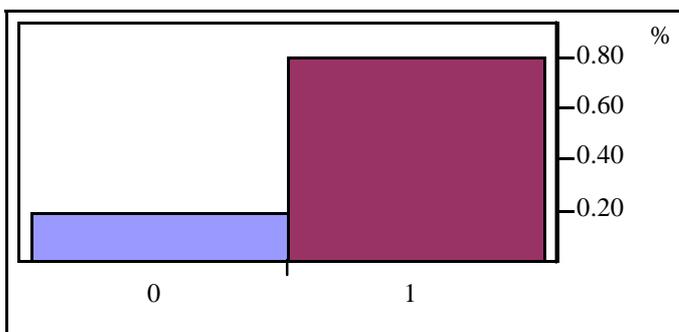


Abb. 6: Gesamtsensitivität der 1. Methode in Prozent. 1= richtig positive Einteilungen, 0= falsch negative Einteilungen

Tab. 9: Gesamtsensitivität und Konfidenzintervalle (CI) der 1.Methode. Gesamtsensitivität und Gesamte falsch negative Einteilungen der 1. Methode in Prozent und Gesamtanzahl (= Gesamt). 1= richtig positive Einteilungen, 0= falsch negative Einteilungen. Total= Gesamtanzahl.

Gesamtsensitivität			
	Gesamt	%	
0	161	19.26	
1	675	80.74	
Total	836	100.00	
Konfidenzintervalle			
	Unteres CI	Oberes CI	1-Alpha
0	16.79	22.07	0.950
1	77.93	83.27	

3.2.1.2 Spezifität

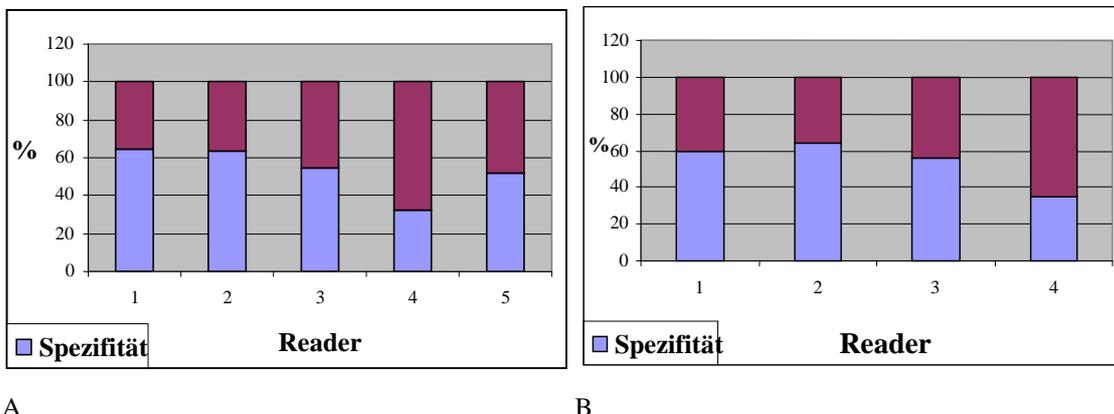


Abb. 7: Spezifität Methode 1. Angaben in Prozent, A: erstes Reading, B: zweites Reading.

Tab. 10: Spezifität und falsch positive Befunde der Methode 1. Angaben in Prozent und Gesamtanzahl der Einteilungen (= Gesamt). 0= richtig negative Einteilungen, 1= falsch positive Einteilungen. Total= Gesamteinteilungen je Untersucher und je Dignität. A: erstes Reading, B: zweites Reading.

Gesamt %	0	1	Total	Gesamt %	0	1	Total
Reader 1	65	36	101	Reader 1	60	41	101
64.36		35.64		59.41		40.59	
Reader 2	64	37	101	Reader 2	65	36	101
63.37		36.63		64.36		35.64	
Reader 3	55	46	101	Reader 3	56	45	101
54.46		45.54		55.45		44.55	
Reader 4	33	68	101	Reader 4	35	66	101
32.67		67.33		34.65		65.35	
Total	217	187	404	Total	216	188	404
Reader 5	11	10	21				
52.38		47.62					
Total	228	197	425				

Abbildung 7 und Tabelle 10 zeigen die Ergebnisse der Spezifität der Methode 1. Wie aus den Tabellen 10 ersichtlich haben 101 der insgesamt 203 ausgewerteten Patientinnen einen gutartigen histologischen Befund. Die Spezifität für Untersucher 1 beträgt 64,36% bei der ersten Befundung bzw. 59,41% bei der zweiten Beurteilung. Dies entspricht einer Gesamtzahl von 65 beziehungsweise 60 Patientinnen, die richtigerweise als gesund eingestuft

Ergebnisse

wurden. Untersucher 2 erreicht im ersten Auswertedurchgang eine Spezifität von 63,37%. Im zweiten Durchgang beträgt sie 64,36%. Es wurden dabei 64 beziehungsweise 65 Patientinnen mit benignen Befunden als gesund kategorisiert. Die Spezifitäten von Untersucher 3 betragen 54,46% und 55,45%. Er hat von den 101 Patientinnen 54 bzw. 55 als richtig gesund eingestuft. Untersucher 4 erreicht eine Spezifität von 32,67% bei der ersten Befundung und 43,65% in der Zweiten. Dies entspricht einer Anzahl von 33 bzw. 35 Patientinnen.

Tab. 11: Signifikanztests der Spezifitätsberechnung der 1. Methode. A: erstes Reading, B: zweites Reading, W= Wahrscheinlichkeit, $W > \text{Chi}^2 =$ Signifikanzniveau.

A			B		
Test	Chi2	Prob>Chi2	Test	Chi2	Prob>Chi2
Likelihood Ratio	26.723	<0.0001	Likelihood Ratio	20.969	<0.0001
Pearson	26.402	<0.0001	Pearson	20.773	<0.0001

Tabelle 11 zeigt die Signifikanztests der Spezifität für die 1. bzw. das 2. Befundung. Wie aus der Tabelle ersichtlich, ist mit beiden Tests eine Signifikanz bewiesen. Dies bedeutet, dass die Spezifität sich von Untersucher zu Untersucher signifikant unterscheidet und dass der Untersucher somit individuell Einfluss auf die Ergebnisse der Spezifität hat.

Aus Abbildung 8 und Tabelle 12 ist die Gesamtspezifität aller Untersucher und jeweils beiden Auswertungen mit der Methode 1 erkennbar. Sie beträgt 53,56%.

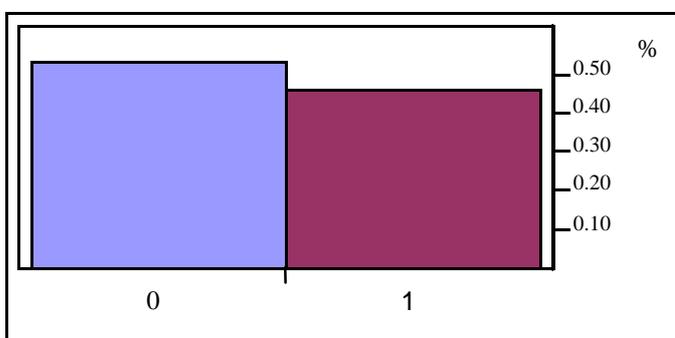


Abb. 8: Gesamtspezifität der 1. Methode in Prozent. 0= richtig negative Einteilungen, 1= falsch positive Einteilungen.

Ergebnisse

Tab. 12: Gesamtspezifität und Konfidenzintervalle (CI) der 1.Methode. Gesamtspezifität und Gesamte falsch positive Einteilungen der 1. Methode in Prozent und Gesamtanzahl (= Gesamt). 0 = richtig negative Einteilungen, 1= falsch positive Einteilungen. Total= Gesamtanzahl.

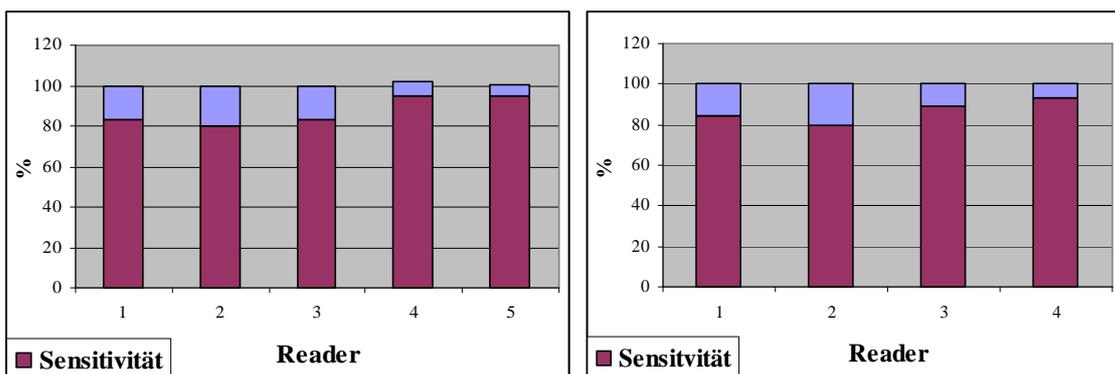
Gesamtspezifität			
	Gesamt	%	
0	444	53.56	
1	385	46,44	
Total	829	100.00	

Konfidenzintervalle			
	Unteres CI	Oberes CI	1-Alpha
0	0.50	0.57	0.950
1	0.43	0.50	

3.2.2 Methode 2: Auswertung nach der subjektiven Beurteilung

3.2.2.1 Sensitivität

Abbildung 9 und Tabelle 13 zeigen die Ergebnisse der Sensitivität der 2. Methode, der subjektiven Beurteilung.



A

B

Abb. 9: Sensitivität Methode 2. Angaben in Prozent, A: erstes Reading, B: zweites Reading.

Ergebnisse

Tab. 13: Sensitivität und falsch negative Befunde der Methode 2. Angaben in Prozent und Gesamtanzahl der Einteilungen (= Gesamt). 0= falsch negative Einteilungen, 1= richtig positive Einteilungen. Total= Gesamteinteilungen je Untersucher und je Dignität. A: erstes Reading, B: zweites Reading.

A				B			
Gesamt %	0	1	Total	Gesamt %	0	1	Total
Reader 1	17 16.67	85 83.33	102	Reader 1	16 15.69	86 84.31	102
Reader 2	20 19.61	82 80.39	102	Reader 2	21 20.59	81 79.41	102
Reader 3	17 16.61	85 83.33	102	Reader 3	11 10.78	91 89.22	102
Reader 4	5 4.90	97 95.10	102	Reader 4	7 6.86	95 93.14	102
Total	59	349	408	Total	55	353	408
Reader 5	1 5.00	19 95.00	20				
Total	60	368	428				

Untersucher 1 erreicht eine Sensitivität von 83,33% bei der ersten Auswertung und in der Zweiten 84,31%. Von den insgesamt 102 Patientinnen mit histologisch malignem Ergebnis wurden von Untersucher 1 85 bzw. 86 anhand der Mammographie in die BI-RADS Kategorie 4 und höher eingeteilt.

Bei Untersucher 2 sind es 82 beim ersten und 81 beim zweiten Auswerten. Seine Sensitivität beträgt 80,39% bzw. 79,41%.

Untersucher 3 teilt 85 Patientinnen und bei der zweiten Beurteilungsrunde 91 Patientinnen in eine ‚maligne‘ Kategorie ein. Die Sensitivitäten betragen daher 83,33% und 89,22%. Bei Untersucher 4 sind es 95,1% und 93,14%. Bei ihm sind es fünf und sieben Patientinnen, die trotz einer malignen Histologie in die BI-RADS Kategorie 3 und niedriger eingeteilt wurden.

Tabelle 14 zeigt die Signifikanztests für die Sensitivität der 2. Methode. Sowohl für die 1. Befundung wie auch für die 2. Befundung wurde mit dem Likelihood-Ratio-Test und mit dem Pearson-Test eine Signifikanz bewiesen. Auch mit dieser Methode hat der Untersucher also Einfluss auf die Sensitivität und die Untersucher unterscheiden sich signifikant voneinander.

Ergebnisse

Tab. 14: Signifikanztests der Sensitivitätsberechnung bei der 2. Methode. A: erstes Reading, B: zweites Reading, W= Wahrscheinlichkeit, $W > \text{Chi}^2 =$ Signifikanzniveau.

A			B		
Test	Chi2	Prob>Chi2	Test	Chi2	Prob>Chi2
Likelihood Ratio	14.300	0.0064	Likelihood Ratio	9.534	0.0230
Pearson	12.213	0.0158	Pearson	9.310	0.0254

Abbildung 10 und Tabelle 15 zeigt die Gesamtsensitivität für die Methode 2, also die subjektive Beurteilung der Mammographien. Sie beträgt 86, 24%.

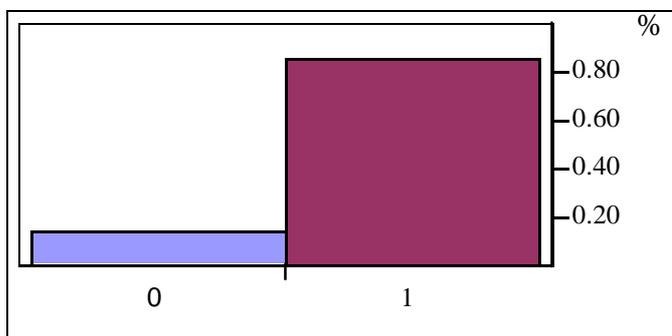


Abb. 10: Gesamtsensitivität der 2. Methode in Prozent. 1= richtig positive Einteilungen, 0= falsch negative Einteilungen.

Tab. 15: Gesamtsensitivität und Konfidenzintervalle (CI) der 1.Methode. Gesamtsensitivität und Gesamte falsch negative Einteilungen der 1. Methode in Prozent und Gesamtanzahl (= Gesamt). 1= richtig positive Einteilungen, 0= falsch negative Einteilungen. Total= Gesamtanzahl.

Gesamtsensitivität			
	Gesamt	%	
0	115	13.76	
1	721	86.24	
Total	836	100.00	
Konfidenzintervalle			
	Unteres CI	Oberes CI	1-Alpha
0	0.12	0.16	0.950
1	0.84	0.88	

3.2.2.2 Spezifität

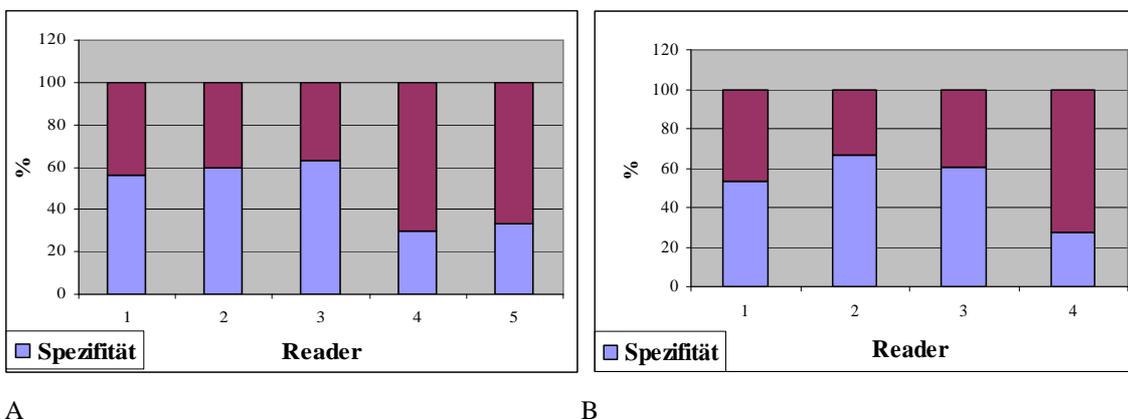


Abb. 11: Spezifität Methode 2. Angaben in Prozent, A: erstes Reading, B: zweites Reading.

Wie schon unter Methode 1 erläutert, haben 101 der insgesamt 203 Patientinnen eine gutartige histologische Diagnose. Die Spezifität der zweiten Methode, also der subjektiven Beurteilung, ergab wie aus Abbildung 11 und Tabelle 16 ersichtlich folgende Ergebnisse:

Untersucher 1 teilte bei der ersten Befundung 57 Patientinnen und bei der zweiten Befundung 54 Patientinnen in die BI-RADS- Kategorie 3 und niedriger ein. Die Spezifität beträgt 56,44% bzw. 53,47%.

Bei Untersucher 2 sind es 59,41% bzw. 66,34%. Von ihm wurden 60 Mammographien beim ersten Auswerten und 67 beim Zweiten als benigne kategorisiert.

Die Spezifität von Untersucher 3 beträgt in der ersten Auswertung 63,37%. Dies entspricht einer Anzahl von 64 Patientinnen. Bei der zweiten Beurteilung sind es 61 Patientinnen. Hier beträgt die Spezifität 60,40%.

Untersucher 4 erkennt 30 bzw. 28 Patientinnen als gesund. Seine Spezifität beträgt 29,7% und 27,72%.

Ergebnisse

Tab. 16: Spezifität und falsch positive Befunde der Methode2. Angaben in Prozent und Gesamtanzahl der Einteilungen (= Gesamt). 0= richtig negative Einteilungen, 1= falsch positive Einteilungen. Total= Gesamteinteilungen je Untersucher und je Dignität. A: erstes Reading, B: zweites Reading.

A				B			
Gesamt	0	1	Total	Gesamt	0	1	Total
%				%			
Reader 1	57 56.44	44 43.56	101	Reader 1	54 53.47	47 46.53	101
Reader 2	60 59.41	41 40.59	101	Reader 2	67 66.34	34 33.66	101
Reader 3	64 63.37	37 36.6	101	Reader 3	61 60.40	40 39.60	101
Reader 4	30 29.70	71 70.30	101	Reader 4	28 27.72	73 72.28	101
Total	211	193	404	Total	210	194	404
Reader 5	7 33.33	14 66.67	21				
Total	218	207	425				

Die Testung der Signifikanz erbrachte für beide Auswertungen ein signifikantes Ergebnis. In Tabelle 17 kann man erkennen, dass das Ergebnis sowohl mit dem Likelihood-Ratio-Test als auch mit dem Pearson-Test deutlich unter dem angesetzten Signifikanzniveau von 0,05 liegt. Die Spezifitäten der Untersucher unterscheiden sich also signifikant voneinander. Die Untersucher haben auch hier Einfluss auf die Ergebnisse der Spezifität.

Tab. 17: Signifikanztests der Spezifitätsberechnung der 2. Methode. A: erstes Reading, B: zweites Reading, W= Wahrscheinlichkeit, $W > \text{Chi}^2 =$ Signifikanzniveau.

A			B		
Test	Chi ²	Prob>Chi ²	Test	Chi ²	Prob>Chi ²
Likelihood Ratio	36.005	<0.0001	Likelihood Ratio	31.808	<0.0001
Pearson	31.105	<0.0001	Pearson	31.178	<0.0001

Abbildung 12 und Tabelle 18 zeigen das Ergebnis der Gesamtspezifität. Das Gesamtergebnis kommt dadurch zustande, indem man alle 4 Untersucher und beide Auswertungen zusammen betrachtet. Dadurch kommt es zu einer Gesamtsumme von 829 benignen Fällen. Davon wurden 428 korrekt klassifiziert. Die daraus berechnete Spezifität beträgt 51,63%.

Ergebnisse

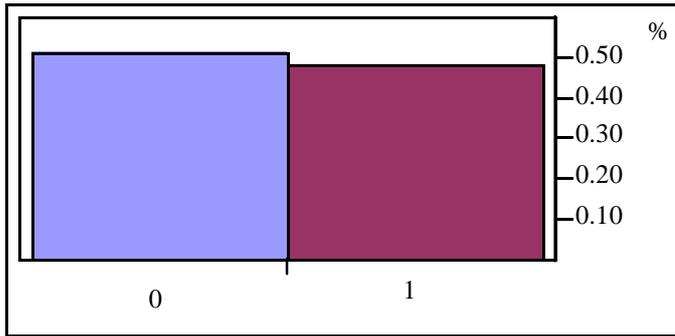


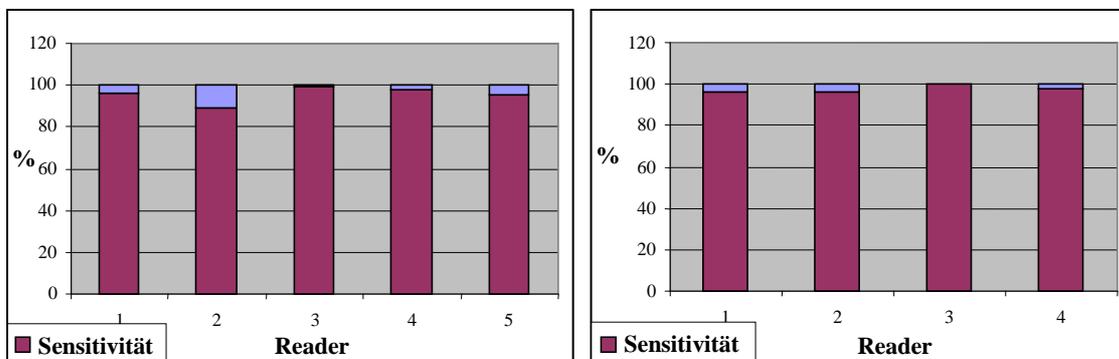
Abb. 12: Gesamtspezifität der 2. Methode in Prozent. 0= richtig negative Einteilungen, 1= falsch positive Einteilungen.

Tab. 18: Gesamtspezifität und Konfidenzintervalle (CI) der 2..Methode. Gesamtspezifität und Gesamte falsch positive Einteilungen der 1. Methode in Prozent und Gesamtanzahl (= Gesamt). 0 = richtig negative Einteilungen, 1= falsch positive Einteilungen. Total= Gesamtanzahl.

Gesamtspezifität			
	Gesamt	%	
0	428	51.63	
1	401	48.37	
Total	829	100.00	
Konfidenzintervalle			
	Unteres CI	Oberes CI	1-Alpha
0	0.48	0.55	0.950
1	0.45	0.52	

3.2.3 Methode 3: Auswertung mit der modifizierten Matrix

3.2.3.1 Sensitivität



A

B

Abb. 13: Sensitivität Methode 3. Angaben in Prozent, A: erstes Reading, B: zweites Reading.

Tab. 19: Sensitivität und falsch negative Befunde der Methode 3. Angaben in Prozent und Gesamtanzahl der Einteilungen (= Gesamt). 0= falsch negative Einteilungen, 1= richtig positive Einteilungen. Total= Gesamteinteilungen je Untersucher und je Dignität. A: erstes Reading, B: zweites Reading.

A				B			
Gesamt %	0	1	Total	Gesamt %	0	1	Total
Reader 1	4 3.92	98 96.08	102	Reader 1	4 3.92	98 96.08	102
Reader 2	11 10.78	91 89.22	102	Reader 2	4 3.92	98 96.08	102
Reader 3	1 0.98	101 99.02	102	Reader 3	0 0.00	102 100.00	102
Reader 4	2 1.96	100 98.04	102	Reader 4	2 1.96	100 98.04	102
Total	18	390	408	Total	10	398	408
Reader 5	1 5.00	19 95.00					
Total	19	409	428				

Die dritte Methode ist die Auswertung mit der modifizierten Matrix. Hiermit ergaben die Ergebnisse der Sensitivität folgende Werte: Untersucher 1 und 4 haben bei der ersten und zweiten Befundung jeweils identische Ergebnisse erzielt. Die Sensitivität von Untersucher 1 beträgt 96,08 %.

Die von Untersucher 4 98,04 %. Von den 102 Patientinnen mit malignen

Ergebnisse

histologischen Diagnosen wurden entsprechend 98 und 100 in eine maligne Kategorie eingeteilt. Die Sensitivität von Untersucher 2 beträgt beim der ersten Auswertung 89,22% und bei der zweiten Auswertung 96,08%

Untersucher 3 erkannte im ersten Durchlauf 101 aller malignen Fälle. Im zweiten Durchlauf waren es alle der insgesamt 102 Patientinnen. Dementsprechend beträgt die Sensitivität 99,02% bzw. 100%.

Abbildungen 13 und Tabelle 19 veranschaulichen die genauen Ergebnisse der Sensitivität.

Tabelle 20 zeigt das Ergebnis der Signifikanztestung. Die Sensitivität der ersten Befundung ist signifikant. Dies trifft für die zweite Befundung nicht zu. Hier unterscheidet sich das Ergebnis der Sensitivität nicht signifikant von Untersucher zu Untersucher.

Tab. 20: Signifikanztests der Sensitivitätsberechnung bei der 2. Methode. A: erstes Reading, B: zweites Reading, W= Wahrscheinlichkeit, $W > \text{Chi}^2 =$ Signifikanzniveau.

A			B		
Test	Chi ²	W>Chi ²	Test	Chi ²	W>Chi ²
Likelihood Ratio	13.118	0.0107	Likelihood Ratio	6.738	0.0807
Pearson	14.113	0.0069	Pearson	4.511	0.2114

Die Auswertung mit der modifizierten Matrix zeigt, wie in Abbildung 15 und Tabelle 21 ersichtlich, eine Gesamtsensitivität von 96,53%.

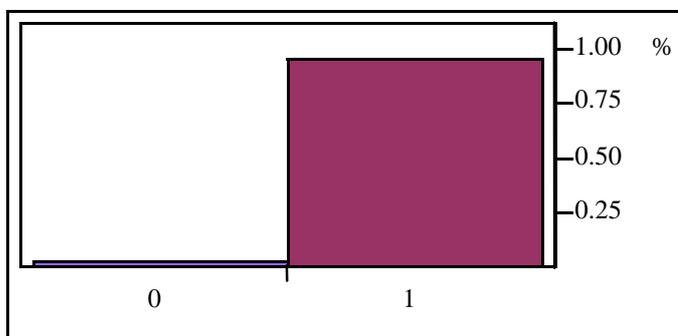


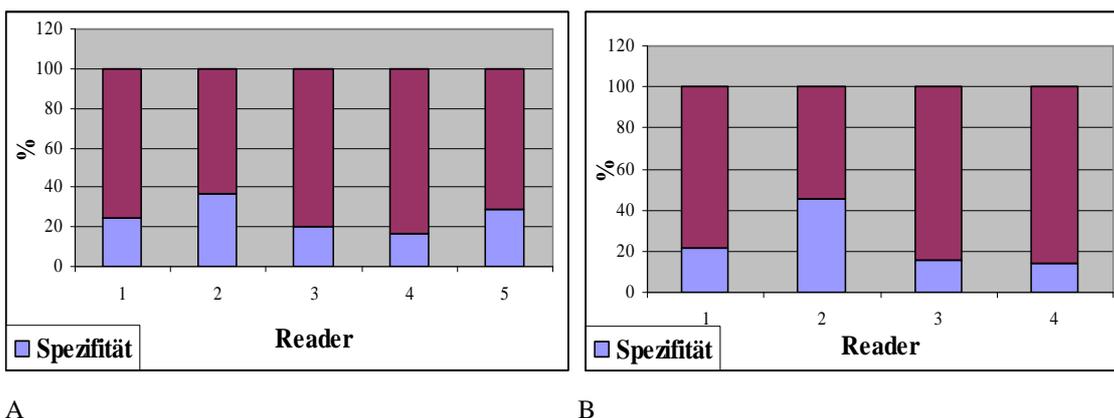
Abb. 14: Gesamtsensitivität der 3. Methode in Prozent. 1= richtig positive Einteilungen, 0= falsch negative Einteilungen.

Ergebnisse

Tab. 21: Gesamtsensitivität und Konfidenzintervalle (CI) der 3..Methode. Gesamtsensitivität und Gesamte falsch negative Einteilungen der 3. Methode in Prozent und Gesamtanzahl (= Gesamt). 1= richtig positive Einteilungen, 0= falsch negative Einteilungen. Total= Gesamtanzahl.

Gesamtsensitivität			
	Gesamt	%	
0	29	3.47	
1	807	96.53	
Total	836	100.00	
Konfidenzintervalle			
	Unteres CI	Oberes CI	1-Alpha
0	2.43	49.38	0.950
1	95.06	97.57	

3.2.3.2 Spezifität



A

B

Abb. 15: Spezifität Methode 3. Angaben in Prozent, A: erstes Reading, B: zweites Reading.

Die Ergebnisse der Spezifität sind in Abbildung 15 und Tabelle 22 dargestellt. Beim der ersten Befundung liegen sie zwischen 16,83% bei Untersucher 4 und 36,63 % bei Untersucher 2.

Bei der zweiten Befundung reichen sie von 13,86% bei Untersucher 4 bis 45,54% bei Untersucher 2.

Untersucher 1 und 3 liegen mit ihren Ergebnissen dazwischen und haben bei der ersten Auswertung ein Ergebnis von 24,75% bzw.19,83%. Bei der zweiten Befundung der Mammographien betragen ihre Spezifitäten 21,78% und 15,84%.

Ergebnisse

Tab. 22: Spezifität und falsch positive Befunde der Methode2. Angaben in Prozent und Gesamtanzahl der Einteilungen (= Gesamt). 0= richtig negative Einteilungen, 1= falsch positive Einteilungen. Total= Gesamteinteilungen je Untersucher und je Dignität. A: erstes Reading, B: zweites Reading.

A				B			
Gesamt	0	1	Total	Gesamt	0	1	Total
%				%			
Reader 1	25 24.75	76 75.25	101	Reader 1	22 21.78	79 78.22	101
Reader 2	37 36.63	64 63.37	101	Reader 2	46 45.54	55 54.46	101
Reader 3	20 19.80	81 80.20	101	Reader 3	16 15.84	85 84.16	101
Reader 4	17 16.83	84 83.17	101	Reader 4	14 13.86	87 86.14	101
Total	99	305	404	Total	98	306	404
Reader 5	6 28.57	15 71.43	21				
Total	105	320	425				

Tabelle 23 zeigt die Ergebnisse der Signifikanztestung. Die beiden Befundungen zeigen ein signifikantes Ergebnis. Auch die Spezifität der dritten Methode unterscheidet sich also signifikant von Untersucher zu Untersucher.

Tab. 23: Signifikanztests der Spezifitätsberechnung der 3. Methode. A: erstes Reading, B: zweites Reading, W= Wahrscheinlichkeit, $W > \text{Chi}^2$ = Signifikanzniveau.

A			B		
Test	Chi ²	W>Chi ²	Test	Chi ²	W>Chi ²
Likelihood Ratio	12.271	0.0154	Likelihood Ratio	32.994	<0.001
Pearson	12.566	0.0136	Pearson	35.081	<0.001

Das Ergebnis der Gesamtspezifität ist in Abbildung 16 und Tabelle 24 zu sehen. Es beträgt 24,49%.

Ergebnisse

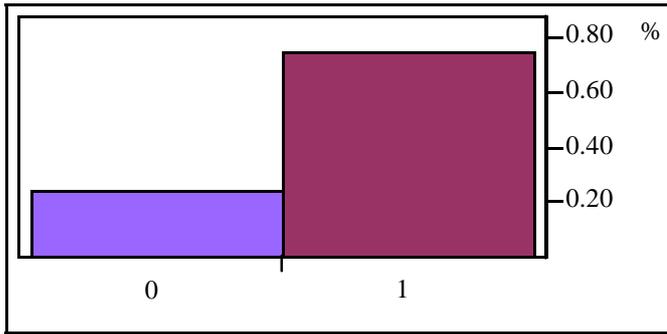


Abb. 16: Gesamtspezifität der 3. Methode in Prozent. 0= richtig negative Einteilungen, 1= falsch positive Einteilungen.

Tab. 24: Gesamtspezifität und Konfidenzintervalle (CI) der 3..Methode. Gesamtspezifität und Gesamte falsch positive Einteilungen der 3. Methode in Prozent und Gesamtanzahl (= Gesamt). 0 = richtig negative Einteilungen, 1= falsch positive Einteilungen. Total= Gesamtanzahl.

Gesamtspezifität			
	Gesamt	%	
0	203	24.49	
1	626	75.51	
Total	829	100.00	
Konfidenzintervalle			
	Unteres CI	Oberes CI	1-Alpha
0	21.68	27.53	0.950
1	72.47	78.32	

3.2.4 Vergleich der Gesamtsensitivitäten und –spezifitäten

Die in Kapitel 3.1.1, 3.1.2 und 3.1.3 aufgeführten Gesamtsensitivitäten und Gesamtspezifitäten werden in Abbildung 17 und 18 noch einmal zur besseren Vergleichbarkeit der Ergebnisse gegenübergestellt.

Bei ihrer Berechnung werden alle Untersucher gemeinsam betrachtet. Jeder Patient kommt also 4-5-mal in der Gesamtanzahl vor. Außerdem werden auch beide Auswertungen gemeinsam betrachtet. Alle Patienten sind also pro Untersucher doppelt vertreten. Die Gesamtsumme der dadurch entstandenen ‚Fälle‘ summiert sich so auf 1667.

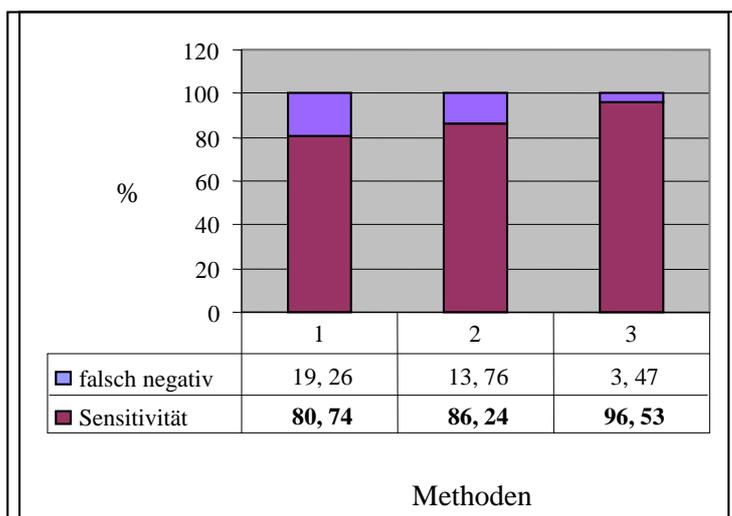


Abb. 17: Vergleich der Gesamtsensitivitäten.

Die Sensitivität von Methode 1, die Auswertung mit der Matrix, hat als Ergebnis 80,74%. Die subjektive Beurteilung der Mammographien, also Methode 2 kommt auf 86,24%. Die dritte Methode, die Auswertung mit der modifizierten Matrix, hat eine Sensitivität von 96,53%.

Ergebnisse

Tab. 25: Signifikanztests der Gesamtsensitivätsberechnung. A: Vergleich Methode 2 mit Methode 1, B: Vergleich Methode 2 mit Methode 3. W= Wahrscheinlichkeit, $W > \text{Chi}^2$ = Signifikanzniveau.

A			B		
Test	Chi ²	W>Chi ²	Test	Chi ²	W>Chi ²
Likelihood Ratio	9.219	0.0024	Likelihood Ratio	59.801	<0.001
Pearson	9.182	0.0024	Pearson	56.201	<0.001

Mit dem Likelihood-Ratio-Test und dem Pearsonsontest wurde die Signifikanz getestet. Tabelle 25 A zeigt die Ergebnisse des Vergleichs von Methode 2 mit Methode 1. Tabelle 25 B den Vergleich von Methode 2 mit Methode 3.

Die Vergleiche der Gesamtsensitivitäten ergaben jeweils ein signifikantes Ergebnis. Die Methoden unterscheiden sich hinsichtlich der Sensitivität also signifikant voneinander. Sie haben somit Einfluss auf die Sensitivität.

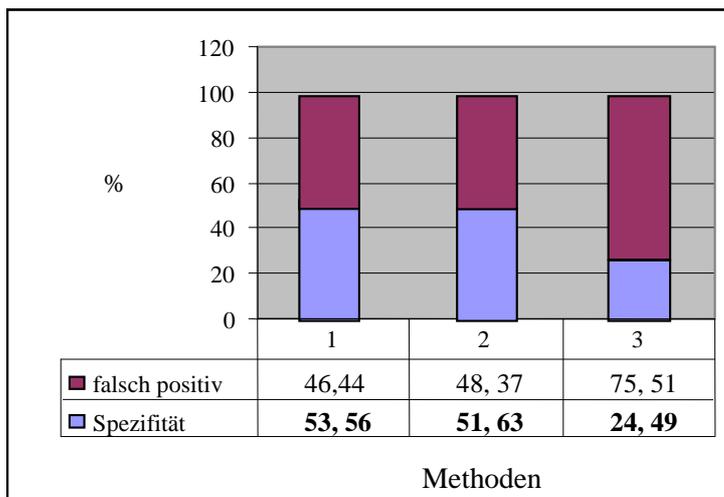


Abb. 18: Vergleich der Gesamtspezifitäten.

Die Ergebnisse der Gesamtspezifitäten sind für Methode 1, die Auswertung mit der Matrix, 53,56%, für Methode 2, die subjektive Beurteilung, 51,63% und für die dritte Methode, die Auswertung mit der modifizierten Matrix, 24,49%.

Ergebnisse

Tab. 26: Signifikanztests der Gesamtspezifitätsberechnung. A: Vergleich Methode 2 mit Methode 1, B: Vergleich Methode 2 mit Methode 3. W= Wahrscheinlichkeit, $W > \text{Chi}^2$ = Signifikanzniveau.

A			B		
Test	Chi ²	W>Chi ²	Test	Chi ²	W>Chi ²
Likelihood Ratio	0.619	0.4313	Likelihood Ratio	131.719	<0.001
Pearson	0.619	0.4313	Pearson	129.524	<0.001

Tabelle 26 A und B zeigen die Ergebnisse der Signifikanztest. Der Vergleich von Methode 2 mit Methode 3 erbrachte ein signifikantes Ergebnis. Beide Methoden unterscheiden sich hinsichtlich der Spezifität signifikant voneinander und haben Einfluss auf die Spezifität.

Beim Vergleich der Methode 2 mit Methode 1 konnte kein signifikantes Ergebnis erzielt werden.

3.2.5 Vergleich der Readings

Das folgende Kapitel zeigt die Ergebnisse des Vergleichs der beiden Auswertungen.

Sie sind jeweils für jeden Untersucher getrennt aufgeführt. Verglichen wurde nur die Häufigkeit der Einteilung in eine ‚benigne‘ bzw. ‚maligne‘ Kategorie, unabhängig von dem tatsächlichen histologischen Ergebnis. Außerdem wurde der Einfluss der verschiedenen Methoden darauf betrachtet. Dabei wurden nur zwei Methoden ausgewertet. Methode 1, die Einteilung nach der Matrix, und Methode 2, die Subjektive Einteilung.

3.2.5.1 Reader 1

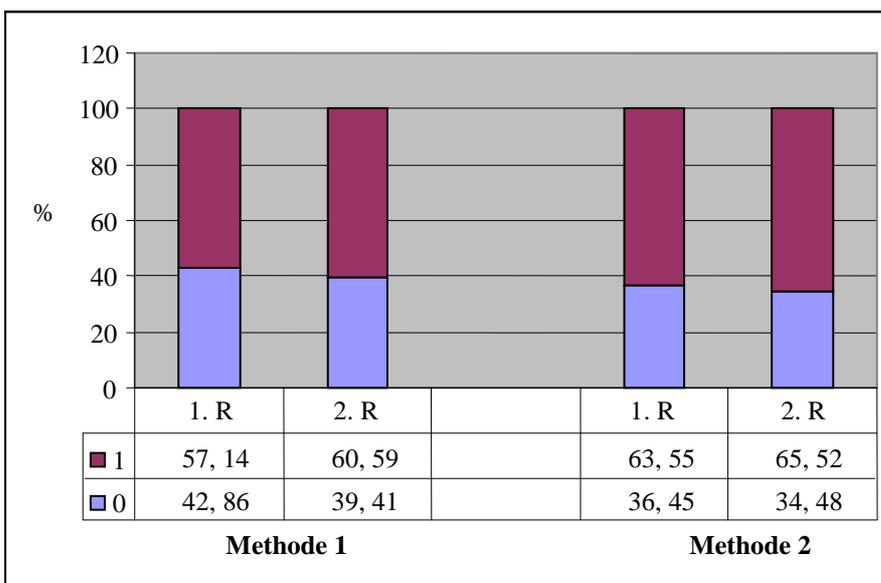


Abb. 19: Vergleich des 1. und 2. Readings bei Untersucher 1. 0= Häufigkeit der Einteilungen in eine benigne Kategorie. 1= Häufigkeit der Einteilungen in eine maligne Kategorie in Prozent. 1.R= 1. Reading. 2.R= 2. Reading.

Abbildung 19 zeigt die Einteilungen für Untersucher 1. Untersucher 1 hat bei der ersten Auswertung und bei der Einteilung mit der Matrix 116 Patienten als maligne klassifiziert, also in BI-RADS Kategorie 4 und höher eingestuft. Bei 87 Patientinnen der Befund als benigne angesehen und in Kategorie 3 und niedriger eingeteilt. Bei der zweiten Auswertung sind es 123 Patientinnen die als maligne und 80 die als benigne angesehen wurden.

Mit der subjektiven Einschätzung, also der zweiten Methode, wurden von Untersucher 1 129 Patientinnen bei der ersten Befundung und 133 Patientinnen bei der zweiten Befundung als maligne angesehen. Entsprechend wurden 74 und 70 Patienten in den beiden Auswertungen als benigne eingestuft.

Die Signifikanztestung mit dem Likelihood Ratio- Test und dem Pearson- Test ergab für die Methode 1 einen p-Wert von 0, 48 und für die Methode einen p-Wert von 0, 68. Die beiden Auswertungen unterscheiden sich also bei beiden Methoden nicht signifikant voneinander.

3.2.5.2 Reader 2

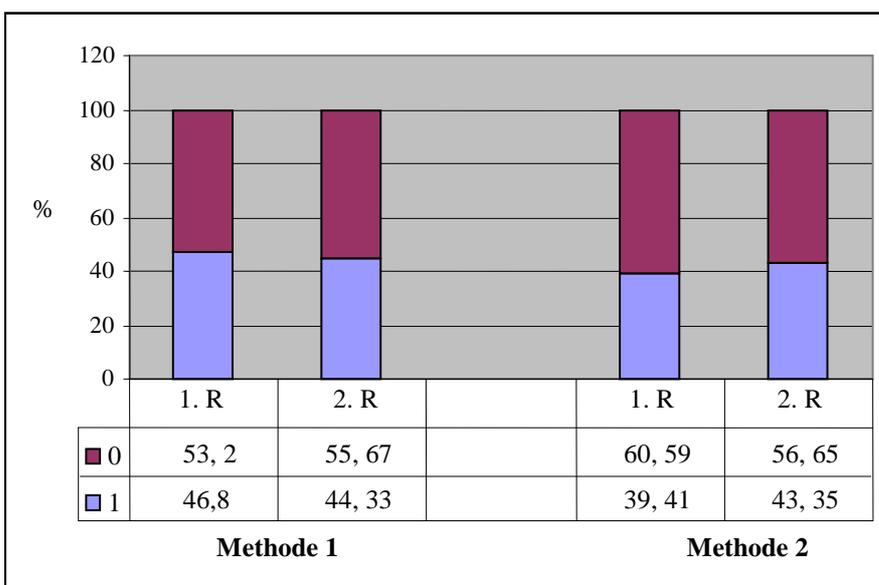


Abb. 20: Vergleich des 1. und 2. Readings bei Untersucher 2. 0= Häufigkeit der Einteilungen in eine benigne Kategorie. 1= Häufigkeit der Einteilungen in eine maligne Kategorie in Prozent. 1.R= 1. Reading. 2.R= 2. Reading.

In Abbildung 20 sind die Ergebnisse des Vergleichs der beiden Befundungen bei Untersucher 2 zu sehen.

Bei Einteilung mit der Matrix sind von Untersucher 2 beim ersten Durchlauf 95 Patientinnen ‚benigne‘ und 108 als ‚maligne‘ klassifiziert worden. Beim zweiten Durchlauf sind es 90 und 113 Patientinnen.

Bei der subjektiven Einschätzung wurde bei der ersten Auswertung bei 123 und

bei der zweiten Auswertung bei 115 Patientinnen der Befund als maligne angesehen. Dementsprechend wurde bei 80 bzw. 88 der Befund als gutartig gesehen.

Auch beim Vergleich der Auswertungen bei Untersucher 2 ergaben die Signifikanztestungen kein signifikantes Ergebnis. Für die Methode 1 lag der p-Wert bei 0,62 für die Methode 2 bei 0,42. Die beiden Auswertungen sind also nicht signifikant unterschiedlich.

3.2.5.3 Reader 3

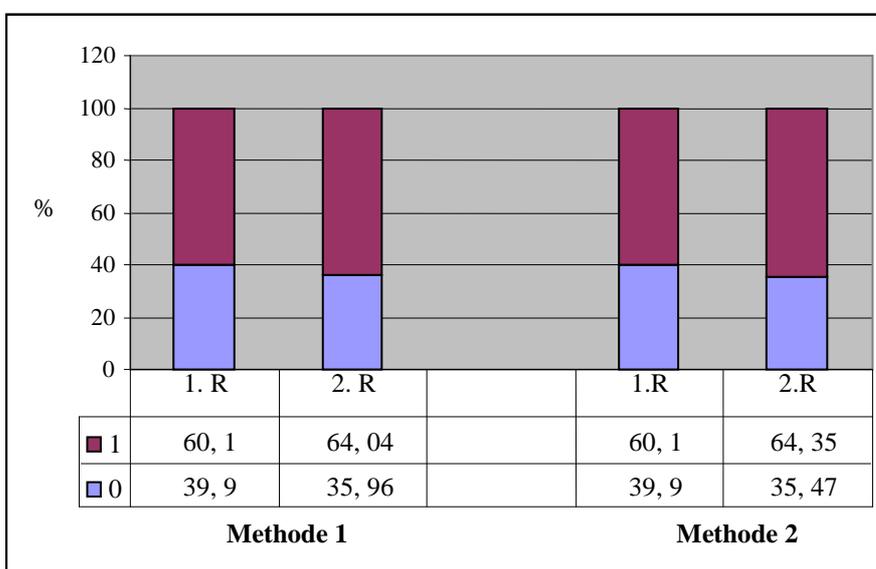


Abb. 21: Vergleich des 1. und 2. Readings bei Untersucher 3. 0= Häufigkeit der Einteilungen in eine benigne Kategorie. 1= Häufigkeit der Einteilungen in eine maligne Kategorie in Prozent. 1.R= 1. Reading. 2.R= 2. Reading.

Untersucher 3 hat wie aus Abbildung 21 ersichtlich mit der Methode 1 122, der insgesamt 203 Mammographien, bei der ersten Befundung in eine maligne Kategorie eingeteilt. Beim der zweiten Befundung waren es 130. Mit der subjektiven Einschätzung sind es 122 beziehungsweise 131.

Dementsprechend wurden mit Methode 1 81 Mammographien beim ersten Auswerten und 73 beim zweiten Auswerten in eine benigne Kategorie eingeteilt. Mit Methode 2 wurden bei 81 beziehungsweise 72 Mammographien der Mikrokalk als benigne befundet.

Die Signifikanztestung ergab kein signifikantes Ergebnis. Auch hier unterscheiden sich die beiden Auswertungsrunden nicht signifikant voneinander. Der p-Wert lag bei 0,41 für die Testung der Methode 1 und bei 0,36 für die Methode 2.

3.2.5.4 Reader 4

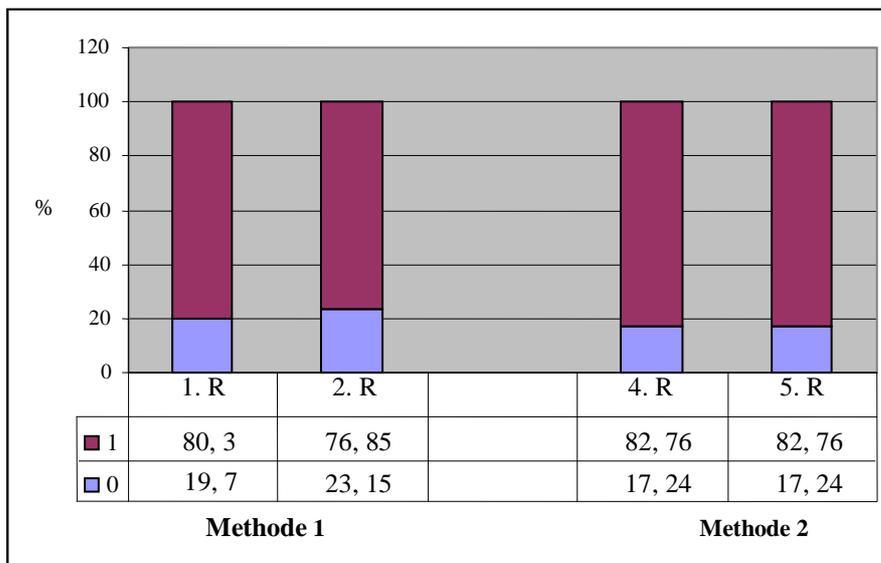


Abb. 22: Vergleich des 1. und 2. Readings bei Untersucher 4. 0= Häufigkeit der Einteilungen in eine benigne Kategorie. 1= Häufigkeit der Einteilungen in eine maligne Kategorie in Prozent. 1.R= 1. Reading. 2.R= 2. Reading.

Untersucher 4 hat mit der Matrix bei der ersten Befundung 163 Patienten als krank klassifiziert. Bei der zweiten Auswertung sind es noch 156. Mit Methode 2 wurden sowohl bei der ersten wie auch bei der zweiten Befundung gleich viele Patienten eingeteilt: 186 ‚Kranke‘ und 35 ‚Gesunde‘.

Die Ergebnisse in Prozent sind in Abbildung 22 zu sehen.

Die Ergebnisse der Signifikanztestung ergaben keine signifikanten Unterschiede zwischen den beiden Methoden. Für Methode 1 lag der p-Wert bei 0,4. Da die Befundungsrunden bei Methode 2 identisch sind, lag der p-Wert hier bei 1,0.

3.2.6 Übereinstimmungen

Im folgenden Kapitel sind die Übereinstimmungen als Gesamtwerte und aufgedgliedert nach Befundung und Untersucher angegeben. Mit Übereinstimmung ist gemeint, wie oft von den Readern die subjektive Einschätzung oder die Einteilung mit der Matrix mit der richtigen histologischen Diagnose -maligne oder benigne- übereinstimmt. Wie in Kapitel 2 erläutert, wurde vereinbart, dass BI-RADS- Kategorie 4 und höher als maligne angesehen wird. Kategorie 3 und niedriger hingegen als benigne.

Dabei kodieren die Zahlen eins und null in den folgenden Tabellen und Diagrammen für Übereinstimmung bzw. für Nichtübereinstimmung:

1= Getroffene Einteilung stimmt mit der Histologie überein.

0= Getroffene Einteilung stimmt **nicht** mit der Histologie überein.

3.2.6.1 Methode 1

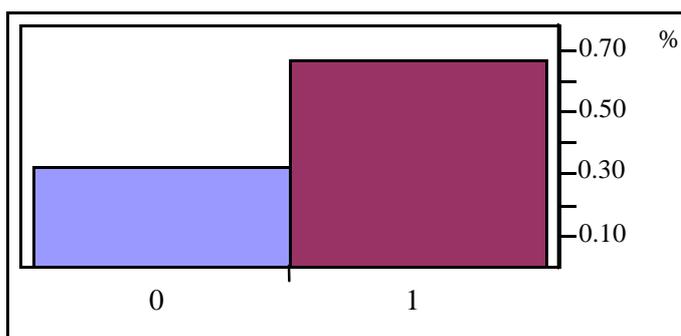


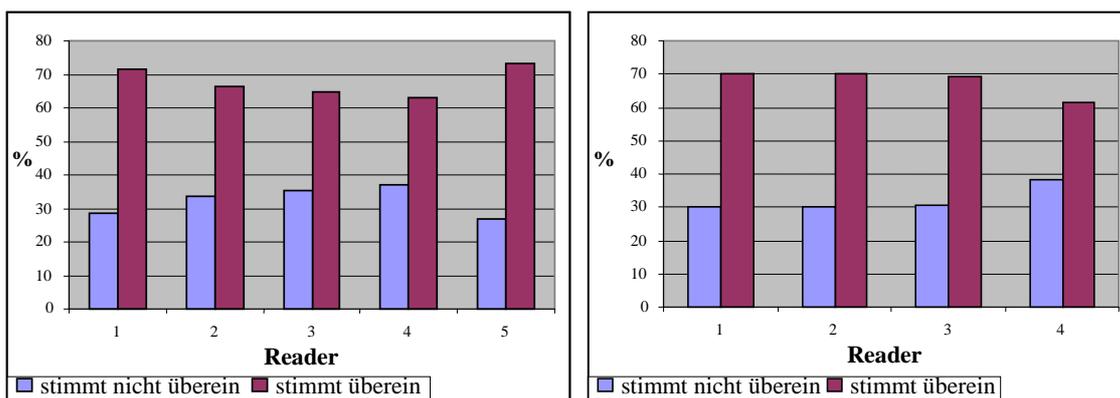
Abb. 23: Gesamtübereinstimmungen der 1. Methode in Prozent. 0= keine Übereinstimmung, 1= Übereinstimmung.

Ergebnisse

Tab. 27: Gesamtübereinstimmungen und Konfidenzintervalle (CI) der 1..Methode. Gesamtübereinstimmungen und Nichtübereinstimmungen der 1. Methode in Prozent und Gesamtanzahl (= Gesamt). 0 = keine Übereinstimmung, 1= Übereinstimmung. Total= Gesamtanzahl.

Gesamtübereinstimmungen			
	Gesamt	%	
0	546	32.79	
1	1119	67.21	
Total	1665	100.00	
Konfidenzintervalle			
	Unteres CI	Oberes CI	1-Alpha
0	30.58	35.09	0.950
1	64.92	69.42	

Abbildung 23 und Tabelle 27 geben die Gesamtübereinstimmungen mit der Auswertung durch die Matrix an. Wie bei vorherigen Gesamtdarstellungen werden auch hier beide Befundungsrunden und alle Untersucher zusammen betrachtet. Dies ergibt eine Gesamtmenge von 1665 Fällen. Bei 1119 davon stimmte die Histologische Diagnose mit der zugeteilten BI-RADS- Kategorie überein.



A B
Abb. 24: Gesamtübereinstimmungen je Untersucher Methode 1. Angaben in Prozent, A: erstes Reading, B: zweites Reading.

Abbildung 24 und Tabelle 28 geben die Ergebnisse für jeden Untersucher einzeln an.

Ergebnisse

Tab. 28: Gesamtübereinstimmungen der Methode 1. Angaben in Prozent und Gesamtanzahl der Einteilungen (= Gesamt). 0 = keine Übereinstimmung, 1= Übereinstimmung. Total= Gesamteinteilungen je Untersucher und je Dignität. A: erstes Reading, B: zweites Reading.

A				B			
Gesamt	0	1	Total	Gesamt	0	1	Total
%				%			
Reader 1	58 28.57	145 71.43	203	Reader 1	61 30.05	142 69.95	203
Reader 2	68 33.50	135 66.50	203	Reader 2	61 30.05	142 69.95	203
Reader 3	72 35.47	131 64.53	203	Reader 3	62 30.54	141 69.46	203
Reader 4	75 36.95	128 63.05	203	Reader 4	78 38.42	125 61.58	203
Total	273	539	812	Total	262	550	812
Reader 5	11 26.83	30 73.17	41				
Total	284	569	853				

Bei der ersten Befundung wurden folgende Ergebnisse erzielt:

Untersucher 1 hat mit Methode 1 bei 145 der 203 Patientinnen die zur histologischen Diagnose passende BI-RADS- Kategorie vergeben. Bei Untersucher 2 sind es 135 richtige Zuteilungen. Untersucher 3 klassifizierte 131 Mammographien in die richtige benigne bzw. maligne Kategorie. Die zur Diagnose passende Kategorie haben 128 Patientinnen bei Untersucher 4.

Bei der zweiten Befundung sind bei Untersucher 1 und 2 142 Patientinnen mit zur Diagnose übereinstimmender Kategorie eingeteilt worden. Bei Untersucher 3 gibt es 141 Übereinstimmungen, bei Untersucher 4 125 Übereinstimmungen.

Tab. 29: Signifikanztests der Berechnung der Gesamtübereinstimmungen der 1. Methode. A: erstes Reading, B: zweites Reading. W= Wahrscheinlichkeit, $W > \text{Chi}^2$ = Signifikanzniveau.

A			B		
Test	Chi^2	$W > \text{Chi}^2$	Test	Chi^2	$W > \text{Chi}^2$
Likelihood Ratio	4.523	0.3399	Likelihood Ratio	4.622	0.2016
Pearson	4.465	0.3467	Pearson	4.711	0.1942

Für Auswertung 1 und 2 ergaben jeweils beide Tests, deren Ergebnisse in

Tabelle 29 zu sehen sind, kein signifikantes Ergebnis. Die Übereinstimmungen unterscheiden sich also nicht signifikant von Untersucher zu Untersucher. Der Untersucher hat keinen Einfluss auf die Übereinstimmungen.

3.2.6.2 Methode 2

Die Gesamtübereinstimmungen bei subjektiver Einteilung der Mammographien sind in Abbildung 25 und Tabelle 30 wiedergegeben. Hier sind von den insgesamt 1665 Fällen bei 1149 einer der histologischen Diagnose entsprechende Kategorie vergeben worden.

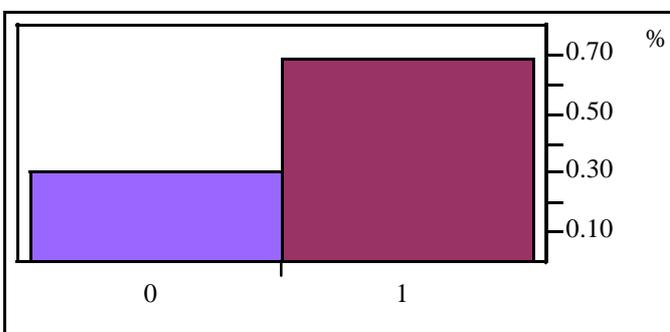


Abb. 25: Gesamtübereinstimmungen der 2. Methode in Prozent. 0= keine Übereinstimmung, 1= Übereinstimmung.

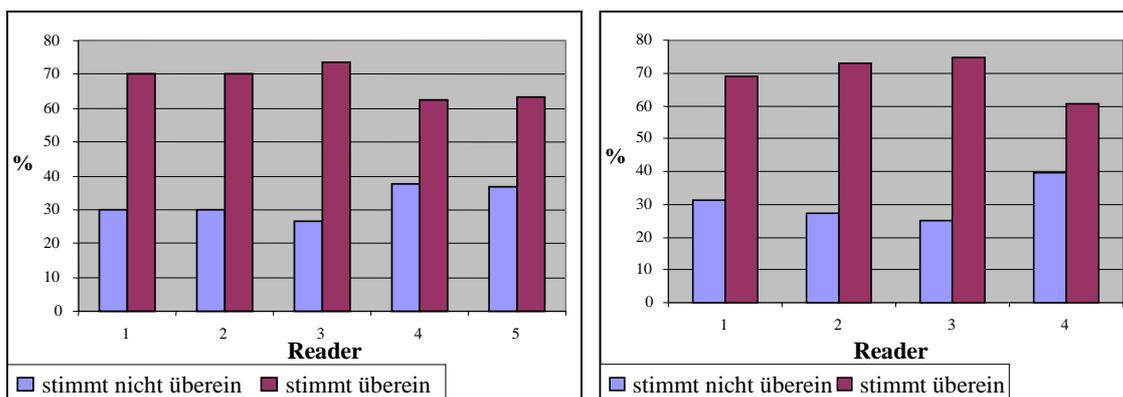
Tab. 30: Gesamtübereinstimmungen und Konfidenzintervalle (CI) der 2..Methode. Gesamtübereinstimmungen und Nichtübereinstimmungen der 2. Methode in Prozent und Gesamtanzahl (= Gesamt). 0 = keine Übereinstimmung, 1= Übereinstimmung. Total= Gesamtanzahl.

Gesamtübereinstimmungen			
	Gesamt	%	
0	516	30.91	
1	1149	69.09	
Total	1665	100.00	
Konfidenzintervalle			
	Unteres CI	Oberes CI	1-Alpha
0	28.16	33.25	0.950
1	66.46	71.28	

Tabelle 31 und Abbildung 26 zeigen die Einzelergebnisse für die Untersucher

Ergebnisse

und die Auswertungsrunden. Untersucher 1 hat in der ersten Auswertung 142 Übereinstimmungen, im Zweiten 140. Bei Untersucher 2 sind bei der ersten Befundung ebenfalls 142, in der zweiten jedoch 148. Die richtige Kategorie hat Untersucher 3 bei 149 bzw. 152 Patientinnen vergeben. Bei Untersucher 4 gibt es 127 bzw. 123 Übereinstimmungen.



A

B

Abb. 26: Gesamtübereinstimmungen je Untersucher Methode 2. Angaben in Prozent, A: erstes Reading, B: zweites Reading.

Tab. 31: Gesamtübereinstimmungen der Methode 2. Angaben in Prozent und Gesamtanzahl der Einteilungen (= Gesamt). 0 = keine Übereinstimmung, 1= Übereinstimmung. Total= Gesamteinteilungen je Untersucher und je Dignität. A: erstes Reading, B: zweites Reading.

A

B

Gesamt %	0	1	Total	Gesamt %	0	1	Total
Reader 1	61	142	203	Reader 1	63	140	203
	30.05	69.95			31.03	68.97	
Reader 2	61	142	203	Reader 2	55	148	203
	30.05	69.95			27.09	72.91	
Reader 3	54	149	203	Reader 3	51	152	203
	26.60	73.40			25.12	74.88	
Reader 4	76	127	203	Reader 4	80	123	203
	37.44	62.56			39.41	60.59	
Total	252	560	812	Total	249	563	812
Reader 5	15	26	41				
	36.59	63.41					
Total	267	586	853				

Ergebnisse

Tab. 32: Signifikanztests der Berechnung der Gesamtübereinstimmungen der 2. Methode. A: erstes Reading 3, B: zweites Reading. W= Wahrscheinlichkeit, $W > \text{Chi}^2$ = Signifikanzniveau.

A			B		
Test	Chi ²	W>Chi ²	Test	Chi ²	W>Chi ²
Likelihood Ratio	6.410	0.1706	Likelihood Ratio	11.281	0.0103
Pearson	6.470	0.1667	Pearson	11.463	0.0095

In Tabelle 32 A ist das Ergebnis für die Signifikanztestung der ersten Auswertung zu sehen. Mit beiden Tests konnte keine Signifikanz bewiesen werden. Hier haben die Untersucher also keinen Einfluss auf die Übereinstimmungen. Bei der zweiten Auswertung, die in Tabelle 32 B zu sehen ist sind die Ergebnisse signifikant. Hier unterscheiden sich die Übereinstimmungen der Untersucher signifikant voneinander. Die Untersucher haben also Einfluss auf die Ergebnisse der Übereinstimmungen.

3.2.6.3 Methode 3

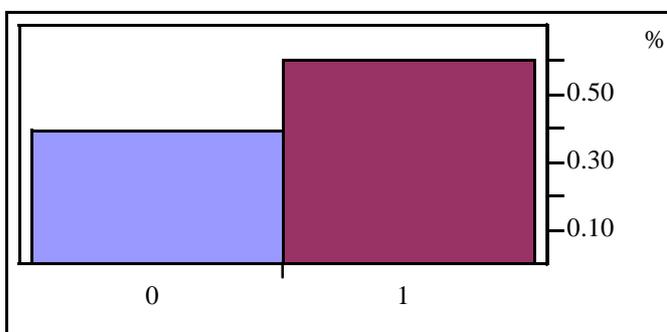


Abb. 27: Gesamtübereinstimmungen der 3. Methode in Prozent. 0= keine Übereinstimmung, 1= Übereinstimmung.

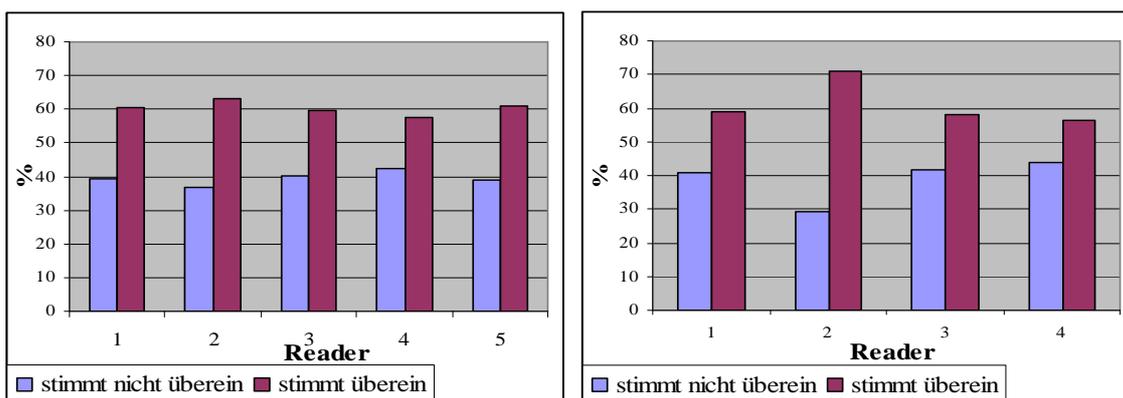
Abbildung 27 und Tabelle 33 zeigen die Ergebnisse der Gesamtübereinstimmungen für die Auswertung mit der modifizierten Matrix. In 60, 66% beziehungsweise in 1010 Fällen stimmt hier die Diagnose mit der vergebenen Kategorie überein.

Ergebnisse

Tab. 33: Gesamtübereinstimmungen und Konfidenzintervalle (CI) der 3..Methode. Gesamtübereinstimmungen und Nichtübereinstimmungen der 3. Methode in Prozent und Gesamtanzahl (= Gesamt). 0 = keine Übereinstimmung, 1= Übereinstimmung. Total= Gesamtanzahl.

Gesamtübereinstimmungen			
	Gesamt	%	
0	655	39.34	
1	1010	60.66	
Total	1665	100.00	
Konfidenzintervalle			
	Unteres CI	Oberes CI	1-Alpha
0	37.02	41.71	0.950
1	58.29	62.98	

In Abbildung 28 und Tabelle 34 sind die Ergebnisse für alle Untersucher getrennt aufgeführt.



A

B

Abb. 28: Gesamtübereinstimmungen je Untersucher Methode 3. Angaben in Prozent, A: erstes Reading, B: zweites Reading.

Bei der ersten Befundung hat Untersucher 1 von den insgesamt 203 Patientinnen bei 123 eine Übereinstimmung. Untersucher 2 hat 128, Untersucher 3 121 und bei Untersucher 4 findet man 117 Übereinstimmungen. Bei der zweiten Befundung fallen die Übereinstimmungen bei Untersucher 1 auf 120. Bei Untersucher 2 steigen sie auf 144. Untersucher 3 und 4 kommen auf 118 und 114 Einteilungen, die zur Diagnose passen.

Ergebnisse

Tab. 34: Gesamtübereinstimmungen der Methode 3. Angaben in Prozent und Gesamtanzahl der Einteilungen (= Gesamt). 0 = keine Übereinstimmung, 1= Übereinstimmung. Total= Gesamteinteilungen je Untersucher und je Dignität. A: erstes Reading, B: zweites Reading.

A				B			
Gesamt %	0	1	Total	Gesamt %	0	1	Total
Reader 1	80 39.41	123 60.59	203	Reader 1	83 40.89	120 59.11	203
Reader 2	75 36.95	128 63.05	203	Reader 2	59 29.06	144 70.94	203
Reader 3	82 40.39	121 59.61	203	Reader 3	85 41.87	118 58.13	203
Reader 4	86 42.36	117 57.64	203	Reader 4	89 43.84	114 56.16	203
Total	323	489	812	Total	316	496	812
Reader 5	16 39.02	25 60.98	41				
Total	339	514	853				

Die Testung der Signifikanz, deren Ergebnisse in Tabelle 35 zu sehen sind, zeigt für die erste Auswertung einen P-Wert über 0,05. Das Ergebnis ist somit nicht signifikant. Beim der zweiten Auswertung liegt er jedoch unter 0,05. Hier ist die Signifikanz bewiesen.

Tab. 35: Signifikanztests der Berechnung der Gesamtübereinstimmungen der 3. Methode. A: erstes Reading 3, B: zweites Reading. W= Wahrscheinlichkeit, $W > \text{Chi}^2$ = Signifikanzniveau.

A			B		
Test	Chi ²	W>Chi ²	Test	Chi ²	W>Chi ²
Likelihood Ratio	1.302	0.8611	Likelihood Ratio	11.734	0.0084
Pearson	1.300	0.8614	Pearson	11.439	0.0096

3.2.7 Ergebnisse nach BI-RADS- Kategorie

In Tabelle 36 und 37 sind die Ergebnisse der einzelnen BI-RADS- Kategorien aufgeführt. Tabelle 36 stellt die Subjektive Einteilung, der Einteilung des Matrix-Systems gegenüber. Die Spalte ‚Gesamt‘ zeigt die Anzahl von Patienten, die von allen Untersuchern und in beiden Readings in die jeweilige Kategorie eingeteilt wurden. Dadurch steigt die Fallzahl auf eine Summe von 1665 Fällen. In der Spalte ‚Maligne‘ wird aufgezeigt, wie viele davon sich jeweils als tatsächlich maligne herausstellten.

Die Spalte ‚Prozentanteil‘ gibt den Anteil der malignen Fälle an der Gesamtmenge in Prozent wieder. So wurden 143 Patienten in die BI-RADS-Kategorie 2 anhand der subjektiven Beurteilung eingeordnet. Davon waren 18 Fälle maligne. Dies heißt also, dass sich bei 12,59% aller Patienten in BI-RADS Kategorie zwei eine Histologie als maligne herausstellte.

Die Auswertung mit dem Matrix-System kam in derselben Kategorie zu einem Ergebnis von insgesamt 144 Patienten. Von diesen 144 Fällen stellten sich 12 als maligne heraus. Der Prozentanteil beträgt hier 8,33%.

Wie man in Tabelle 36 erkennen kann steigt der Anteil der malignen Fälle mit steigender BI-RADS- Kategorie kontinuierlich an. So beträgt der Prozentanteil in Kategorie 3 24,25% bei der Subjektiven Beurteilung und 32,32% mit dem Matrix-System. In Kategorie 4 sind es 52,89% und 48,35%. In Kategorie 5 finden sich die mit 86,02% beziehungsweise 83,84% höchsten Anteile an malignen Diagnosen.

Keine der Mammographien wurde mit der Matrix oder mit der subjektiven Beurteilung in die Kategorien 0,1 oder 6 eingeteilt.

Ergebnisse

Tab. 36: Einteilungen in die BI-RADS Kategorien durch das Matrix-System und die subjektive Einteilung. Gesamt= Anzahl der Einteilungen in diese Kategorie von allen Untersuchern und beiden Readings. Maligne: Anzahl der malignen Diagnosen in dieser Kategorie. Prozentanteil: Anteil der malignen Diagnosen an der Gesamtmenge in Prozent.

BI-RADS Kategorie	SUBJEKTIV			MATRIX-SYSTEM		
	Gesamt	Maligne	Prozentanteil	Gesamt	Maligne	Prozentanteil
2	143	18	12,59 %	144	12	8,33 %
3	400	97	24,25 %	461	149	32,32 %
4	743	393	52,89 %	608	294	48,35 %
5	379	326	86,02 %	452	379	83,84 %

Tabelle 37 stellt die Subjektive Einteilung der Einteilung des modifizierten Matrix-Systems gegenüber. Auch hier steigt die Anzahl der malignen Fälle mit steigender BI-RADS- Kategorie an. Entsprechend der Modifikation (siehe Kapitel Material und Methoden) findet man hier jedoch andere Werte in Kategorie drei und vier.

Tab. 37: Einteilungen in die BI-RADS Kategorien durch das modifizierte Matrix-System und die subjektive Einteilung. Gesamt= Anzahl der Einteilungen in diese Kategorie von allen Untersuchern und beiden Readings. Maligne: Anzahl der malignen Diagnosen in dieser Kategorie. Prozentanteil: Anteil der malignen Diagnosen an der Gesamtmenge in Prozent.

BI-RADS Kategorie	SUBJEKTIV			MATRIX-MODIFIZIERT		
	Gesamt	Maligne	Prozentanteil	Gesamt	Maligne	Prozentanteil
2	143	18	12,59 %	144	12	8,33 %
3	400	97	24,25 %	88	17	19,32 %
4	743	393	52,89 %	981	426	43,42 %
5	379	326	86,02 %	452	379	83,84 %

Insgesamt wurden mit allen drei Methoden die meisten Fälle in die BI-RADS Kategorie 4 eingeteilt. Die wenigsten Fälle findet man in Kategorie 2. Mit der Matrix werden deutlich mehr Fälle in Kategorie 5 eingeteilt. Auffallend ist auch, dass bei mit der modifizierten Matrix weit mehr Fälle in Kategorie 4 eingeteilt wurden als mit den anderen beiden Methoden. Da die Modifikation aber auch ein Wechsel von BI-RADS 3 auf 4 bei Intermediärem und rundem Mikrokalk vorsieht ist dieses Ergebnis zu erwarten. Insgesamt machen die Einteilungen mit der modifizierten Matrix in Kategorie 4 und 5 über 85% aus.

Tabelle 38 zeigt die Einteilungen in die vier Kategorien für jeden Untersucher getrennt. Dabei werden beide Auswertungsrunden zusammen betrachtet. Die Patienten sind also doppelt vertreten, für jeden Untersucher ergibt sich eine Fallzahl von 406.

Wie auch in Tabelle 36 wird auch hier für alle vier Untersucher deutlich, dass Kategorie 3 und 5 mit der Matrix mehr Einteilungen erhält als mit der subjektiven Einteilung. Kategorie 4 erhält durch die Matrix von allen Untersuchern weniger Einteilungen. Untersucher 2 hat im Vergleich mehr Fälle in eine benigne Kategorie, also Kategorie 2 und 3, eingeteilt als die anderen drei Untersucher. Dagegen teilt Untersucher 4 mehr Fälle in eine maligne Kategorie 4 und 5.

Ergebnisse

Tab. 38: Einteilungen in die BI-RADS Kategorien durch das Matrix-System und die subjektive Einteilung für alle Untersucher getrennt.

BI-RADS Kategorie	Reader Subjektiv Einteilung nach BI-RADS Kategorie			
	1	2	3	4
2	18	89	19	10
3	126	78	134	60
4	183	145	158	242
5	79	93	95	92
BI-RADS Kategorie	Reader Einteilung mit der Matrix			
	1	2	3	4
2	28	72	28	10
3	139	113	126	77
4	140	112	145	193
5	99	109	107	126

3.2.8 Verteilung der Malignitätswahrscheinlichkeit über die Matrix

Tabelle 39 zeigt die Verteilung der Malignitätswahrscheinlichkeit über die Matrix. Als Grundlage dieser Tabelle diente der Auswertebogen (siehe Anhang). Der in der Mammographie gefundene gruppierte Mikrokalk wurde nach seiner Verteilung beurteilt und entsprechend als diffus, rund, regional, segmental oder linear bezeichnet. Außerdem wurde der Einzelkalk je nach Form als typisch benigne, als intermediär oder als typisch maligne bezeichnet. Anlehnend an Tabelle 36 und 37 bezeichnet die Spalte ‚Gesamt‘ die gesamte Anzahl an Patienten. Auch hier ergibt sich eine Gesamtfallzahl von 1665, da beide Auswertungsrunden und alle Untersucher zusammen betrachtet werden. Die Spalte ‚Maligne‘ gibt die Anzahl der malignen Fälle an der Gesamtzahl an und in der Spalte ‚Prozentanzahl‘ steht der entsprechende Wert in Prozent.

Tab. 39: Verteilung der Malignitätswahrscheinlichkeit über die Matrix. Gesamt= Gesamteinteilungen. Maligne: Anzahl der malignen Diagnosen in dieser Einteilung. %= Anteil der malignen Diagnosen an der Gesamtmenge in Prozent.

VERTEILUNG	VERKALKUNG								
	Typisch benigne			Intermediär			Typisch maligne		
	Gesamt	Maligne	%	Gesamt	Maligne	%	Gesamt	Maligne	%
Diffus	28	1	4	28	4	14	11	10	91
Rund/oval	54	8	15	154	52	34	115	43	37
Regional	61	3	5	219	80	37	94	67	71
Segmental	48	10	6	388	174	45	328	278	85
Linear/duktal	13	3	23	44	31	70	80	70	88

Die niedrigste Malignitätswahrscheinlichkeit zeigt sich bei typisch benignen Verkalkungen und diffuser Verteilung des Kalkes über die Brust. Sie beträgt 4%. Die mit 91% höchste Malignitätswahrscheinlichkeit fand sich ebenfalls bei diffuser Verkalkung. Der Form des Einzelkalkes war hier jedoch typisch maligne. Sieht man diesen Wert und den Wert bei rund-ovaler Verteilung und typisch benigner Verkalkung als Ausreißer nach oben an, dann kann man eine

Ergebnisse

Zunahme der Malignitätswahrscheinlichkeit von diffuser Verteilung zu linear-duktalem Verteilung erkennen. Auch bei der Form der Einzelverkalkungen sieht man eine deutliche Zunahme an malignen Fällen von typisch benignen über intermediär nach typisch maligne.

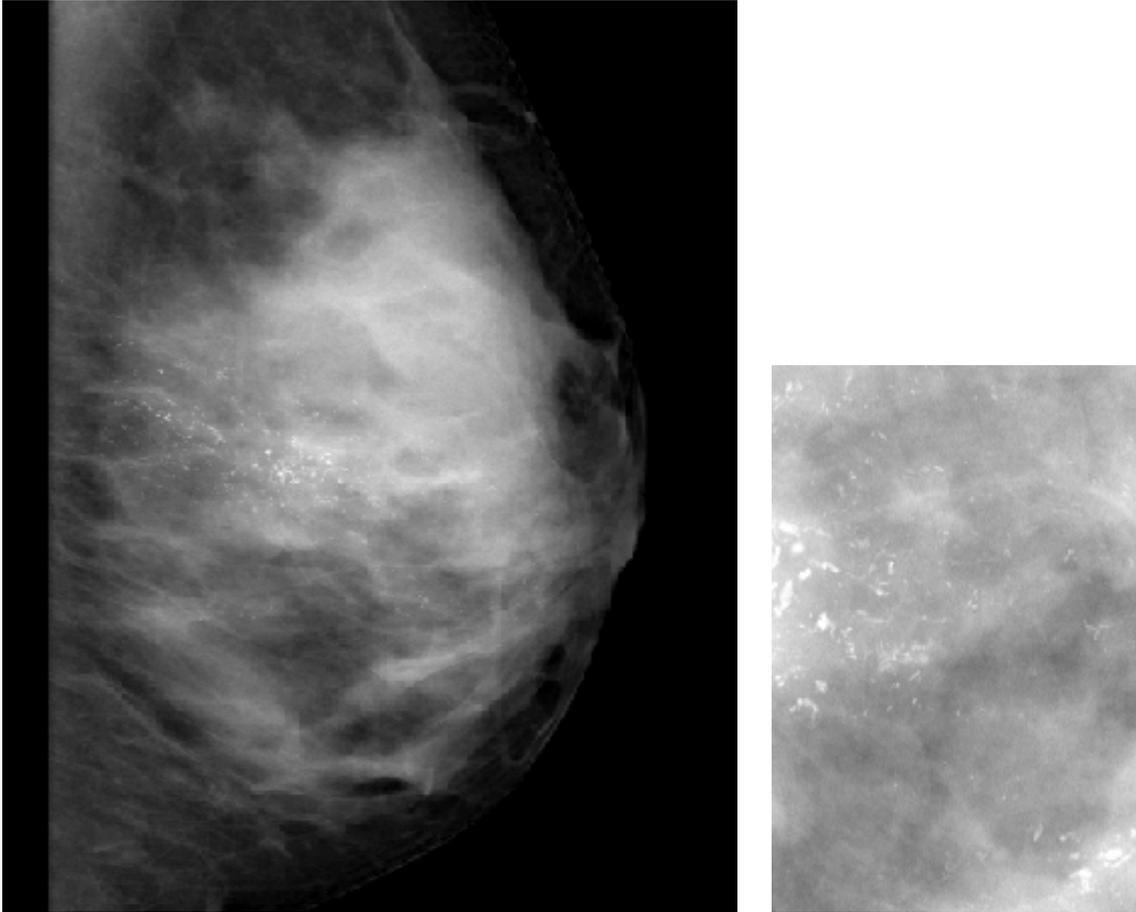
Tabelle 40 stellt die ermittelten Wahrscheinlichkeiten der Matrix, den Wahrscheinlichkeiten der BI-RADS Kategorien gegenüber. Die jeweils zugeordneten BI-RADS Kategorien haben, wie in der Einleitung erläutert, eine bestimmte Malignitätswahrscheinlichkeit. Diese ist in der jeweils mittleren Spalte jeder Verkalkungsart zu finden.

Tab. 40: Gegenüberstellung der Malignitätswahrscheinlichkeiten über die Matrix. Malignitätswahrscheinlichkeiten der BI-RADS Kategorien (= P_M) und der in der Studie ermittelten (P_E) Wahrscheinlichkeiten in Prozent.

VERTEILUNG	VERKALKUNG								
	Typisch benigne			Intermediär			Typisch maligne		
	BI-RADS Kat.	P_M %	P_E %	BI-RADS Kat.	P_M %	P_E %	BI-RADS Kat.	P_M %	P_E %
Diffus	2	<1	4	3	1-3	14	4	20-30	91
Rund/Oval	2	<1	15	3/4	1-3/20-30	34	4	20-30	37
Regional	2	<1	5	3/4	1-3/20-30	37	4	20-30	71
Segmental	3	1-3	6	4	20-30	45	5	>90	85
Linear/Duktal	3	1-3	23	5	20-30	70	5	>90	88

Die in unserer Studie ermittelten Malignitätswahrscheinlichkeiten für die einzelnen BI-RADS Kategorien sind bis auf die segmentale und linear-duktalem Verteilung bei typisch malignem Mikrokalk durchweg höher als die vom ACR veranschlagten Wahrscheinlichkeiten.

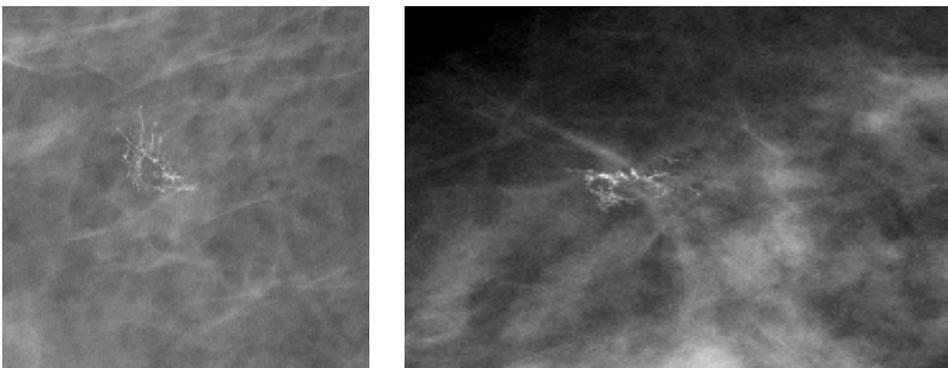
3.3. Beispielmammographien



A

B

Abb. 29: A: MLO- Aufnahme mit typisch segmental gruppiertem Mikrokalk (MK), der auf die Mamille hin zuläuft. B: Ausschnitt des Mikrokalkareals aus der Vergrößerungsaufnahme im CC-Strahlengang mit pleomorphen, typisch-malignen Verkalkungen= BI-RADS 5. Histologie: highgrade DCIS, DIN 3.



A

B

Abb. 30: A: MLO- Strahlengang, Vergrößerungsaufnahme. B: CC- Strahlengang, Vergrößerungsaufnahme. Bizarr geformter, fein-linearer, verzweigender Mikrokalk mit rund-ovalem Verteilungsmuster= BI-RADS 4. Histologie: ‚ungewöhnlicher Befund‘ von im Fettgewebe auskristallisiertem MK.

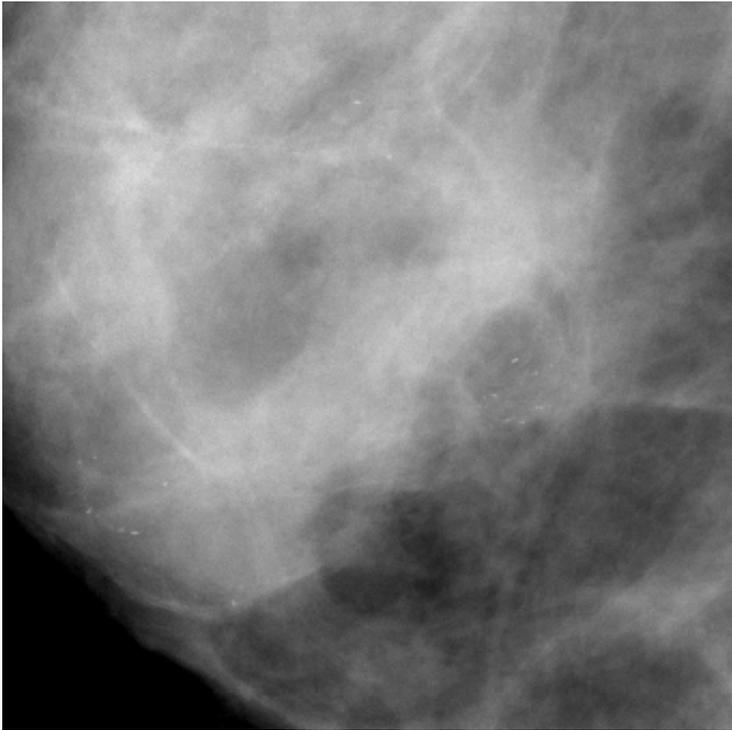


Abb. 31: MLO- Vergrößerungsaufnahme mit granulären, heterogenen Mikroverkalkungen und regionaler Verteilung= BI-RADS 3/4. Dichtes Brustdrüsengewebe bei 44-jähriger Patientin. Histologie: fibrös-zystische Mastopathie.

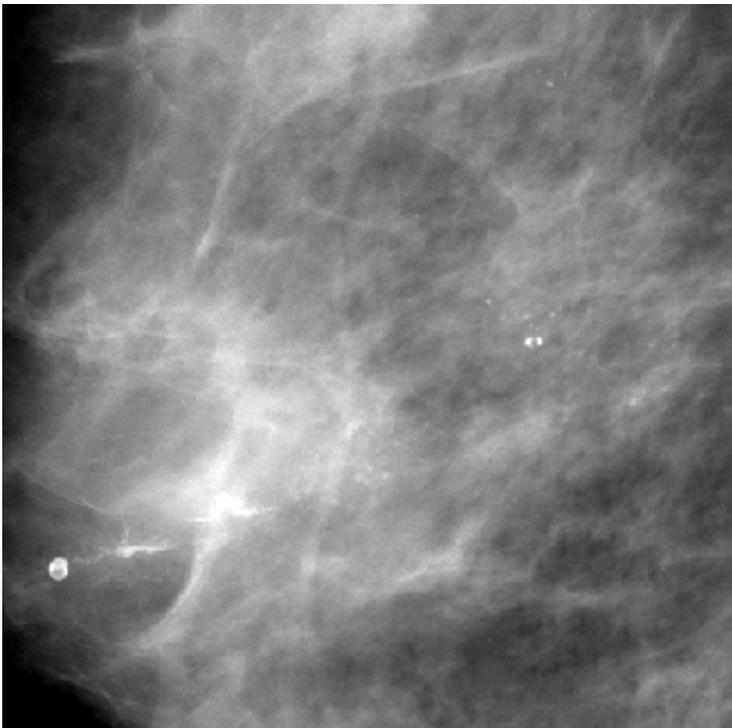
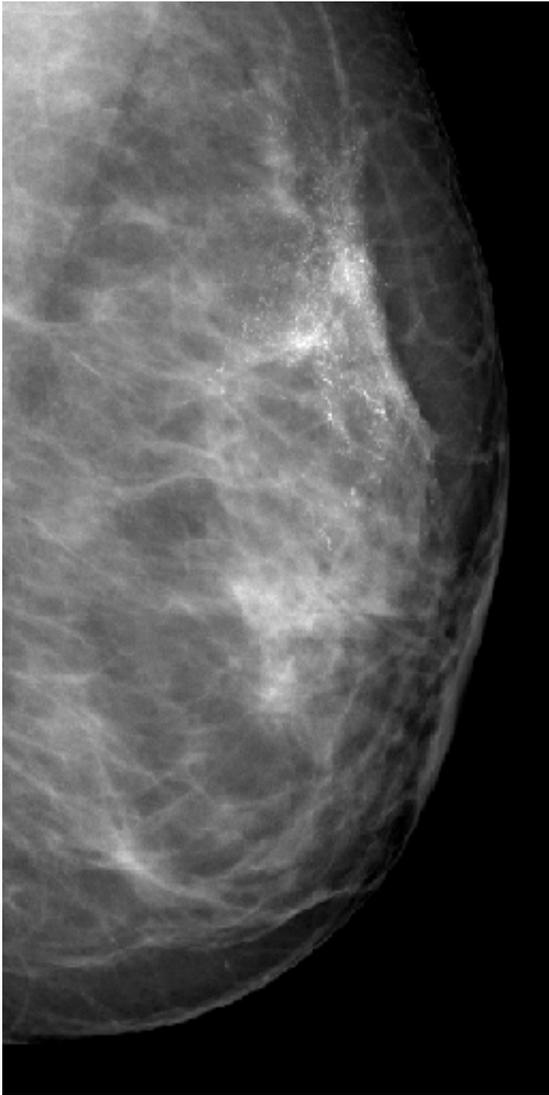
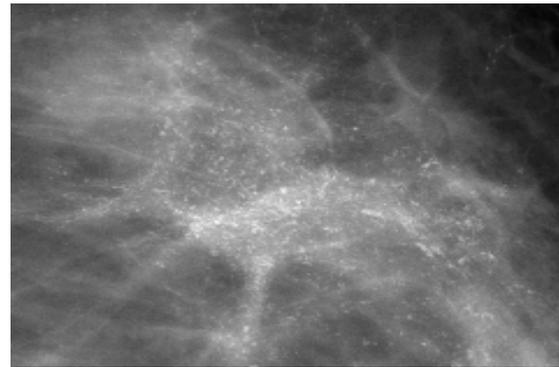


Abb. 32: Vergrößerungsaufnahme im MLO- Strahlengang. Linear- verästelnder, heterogener, teilweise granulärer MK mit segmental- regionalem Verteilungsmuster= BI-RADS 4/5. Außerdem mehrere Makroverkalkungen. Histologie: DCIS, IDC. Makroverkalkungen: Fettgewebsnekrosen.

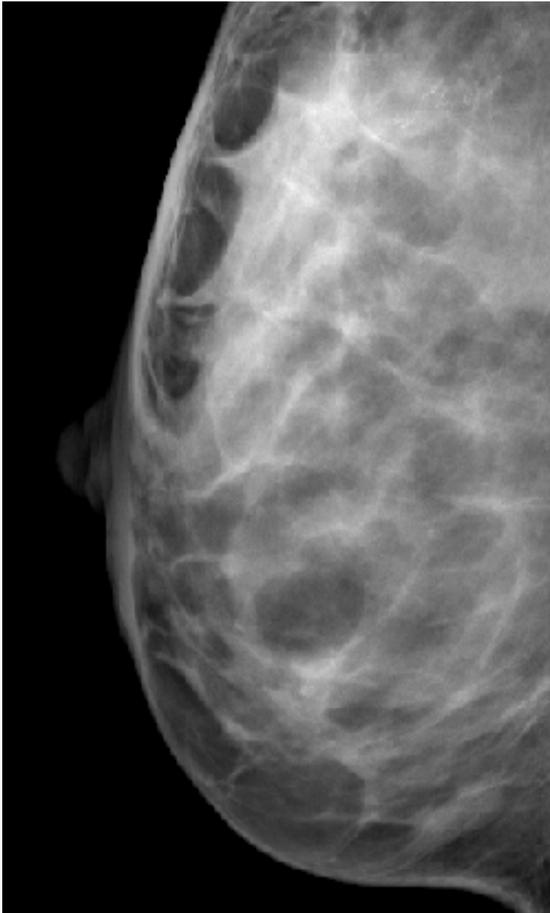


A

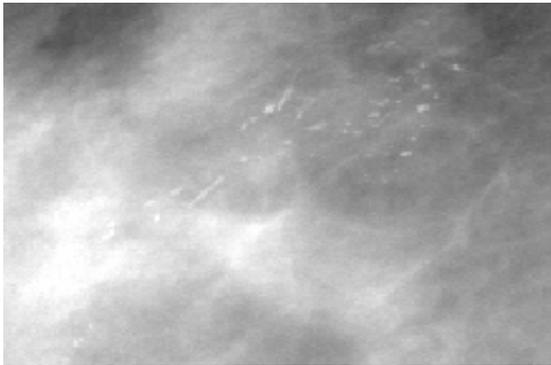


B

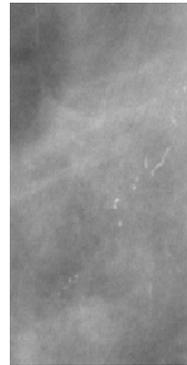
Abb. 33: A: MLO- Aufnahme B: Ausschnitt aus Vergrößerungsaufnahme im CC-Strahlengang. Regional-Segmental angeordneter MK mit typisch maligner Morphologie= BI-RADS 5. Gewebsverdichtung im Areal des MK. Histologie: highgrade DCIS, DIN 3, IDC.



A

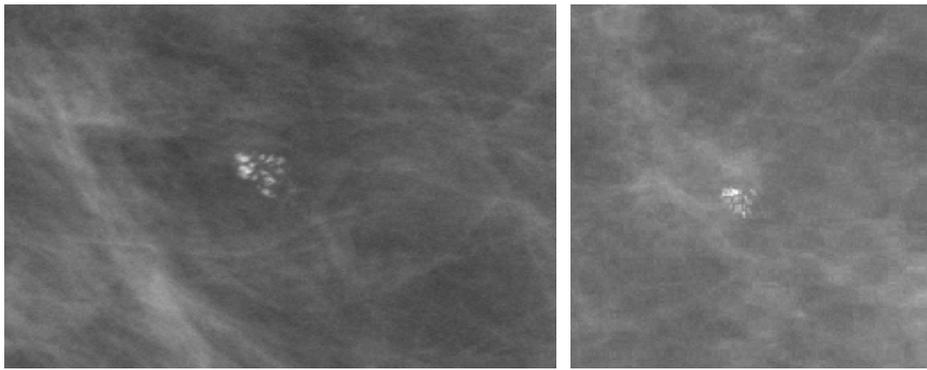


B



C

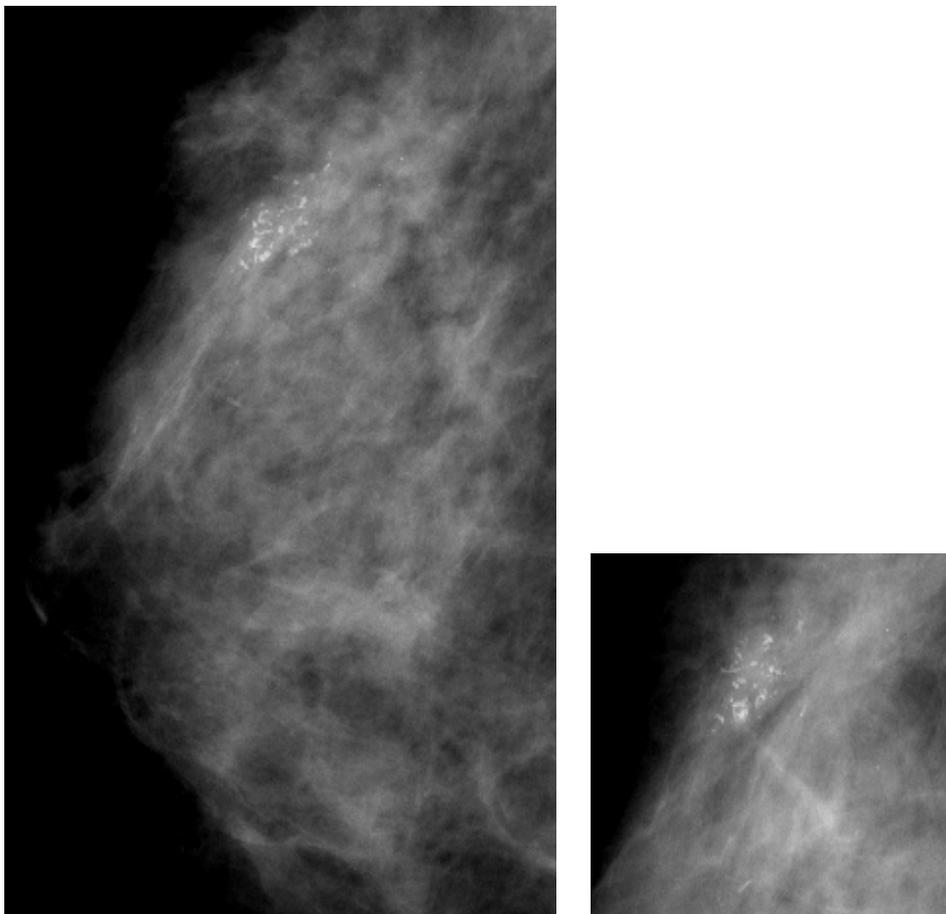
Abb. 34: A: MLO-Aufnahme mit flauem Mikrokalkareal im oberen Quadranten. B: Vergrößerungsaufnahme im MLO- Strahlengang. C: Vergrößerungsaufnahme im CC-Strahlengang. Linearer, pleomorpher MK mit regionaler Verteilung= BI-RADS 4. Histologie: highgrade DCIS, DIN 3.



A

B

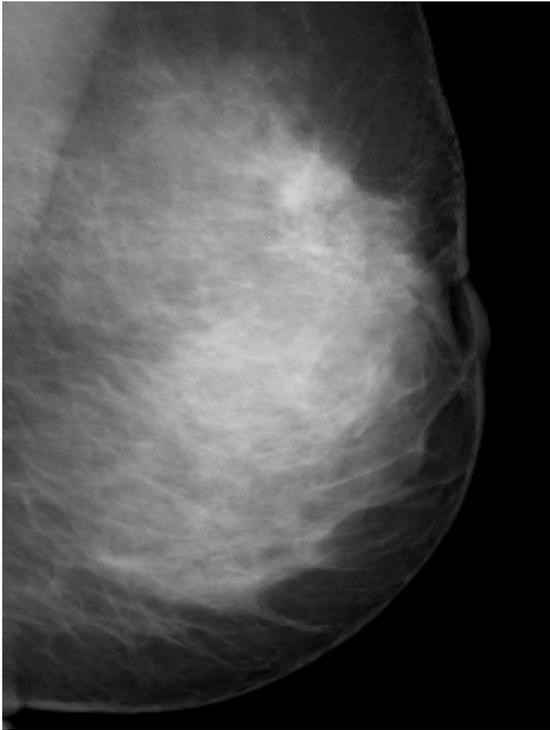
Abb. 35: Ausschnitt aus Vergrößerungsaufnahmen im MLO- (A) und CC- (B) Strahlengang. Rund-ovale Verteilung von granulärem MK= BI-RADS 3/4. Histologie: fibröse Mastopathie.



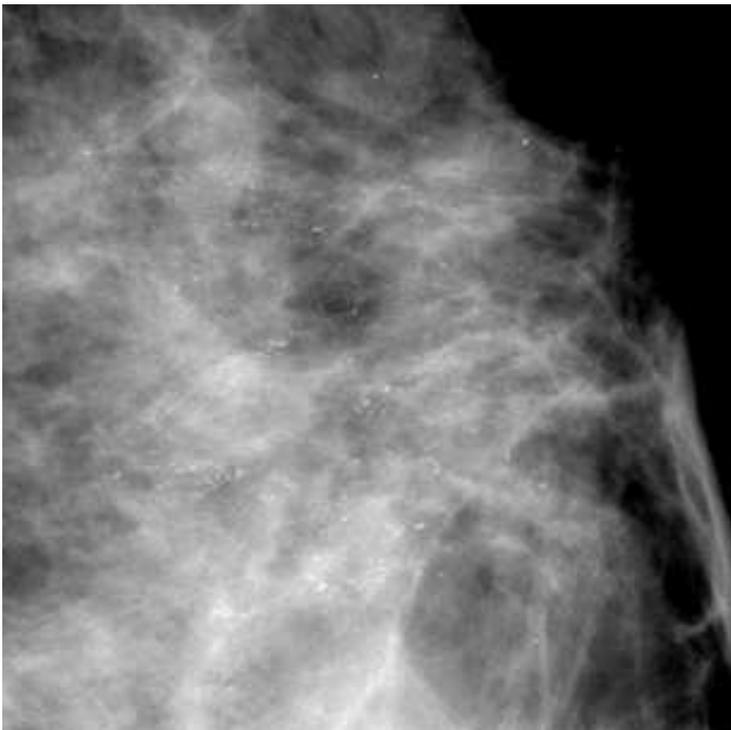
A

B

Abb. 36: A: MLO- Strahlengang B: CC-Strahlengang Regional verteilter MK mit poly- und pleomorpher Morphologie= BI-RADS 4. Histologie: DCIS.



A



B

Abb. 37: A: MLO- Aufnahme mit sternförmigem Herdbefund im oberen Quadranten und Hauteinziehung
B: Vergrößerungsaufnahme im MLO-Strahlengang. Regionale Verteilung von heterogenem MK,
teilweise granulär= BI-RADS 4. Histologie: DCIS, IDC.

4. Diskussion

4.1 Möglichkeiten und Grenzen von BI-RADS

The Breast Imaging Reporting and Data System (BIRADS) des American College of Radiology [70] wurde entwickelt, um die Terminologie bei der Befundung von Mammographien, die Beurteilung der Befunde und die resultierenden diagnostischen und therapeutischen Empfehlungen zu standardisieren [73]. In der Literatur findet man eine Vielzahl von Studien die dieses System evaluieren [79, 91-97].

Neben der Belegung des Nutzens [78, 91, 98-100], heben weitere Arbeiten auch Grenzen beim Einsatz von BI-RADS hervor [28, 78, 86, 91, 97-99, 101-105].

So ist in zahlreichen Untersuchungen eine erhebliche ‚inter- und intraobserver‘ Variabilität beobachtet worden [28, 78, 86, 92, 97]. Trotz Einführung eines standardisierten Lexikons bleibt dieses Problem bei der mammographischen Befundung weiterhin bestehen. Besonders deutlich wird dies bei der Beurteilung und Klassifikation von Mikroverkalkungen, da hierbei die Entscheidung über den Grad der Karzinomverdächtigkeit für die befundenden Radiologen außerordentlich komplex ist [106]. Gülsün et al. kamen zu dem Ergebnis, dass es auch BI-RADS bisher nicht gelingt die Vieldeutigkeit bei der Befundung von Mikroverkalkungen zu reduzieren [95].

Hier kommt besonders zum Tragen, dass keine klaren Empfehlungen bezüglich der Zuordnung der verschiedenen mammographischen Befunde zu den Einteilungskategorien gemacht werden [73, 106].

Es gibt Studien, die sich mit der Klassifikation von Mikroverkalkungen beschäftigen und die aufzeigen, bei welcher Form und Verteilung mit einer höheren Karzinomwahrscheinlichkeit gerechnet werden muss [87, 89, 107, 108]. Jedoch finden sich in der Literatur bisher keine konkreten Zuordnungshilfen zu den BI-RADS Kategorien [81].

Es sind nun aber gerade die Mikroverkalkungen, die durch Ihre Komplexität eine Klassifikation in benigne und maligne Verkalkungen erschweren [109]. Die

Einteilung von anderen mammographischen Befunden, wie Verdichtungen oder Herdbefunden, ist hinsichtlich ihrer Verdächtigkeit verlässlicher [110].

Mit der Matrix ist erstmals eine Interpretationshilfe für Mikroverkalkungen entworfen worden, die je nach Form und Verteilung der Verkalkungen eine konkrete Empfehlung für eine Einteilung in eine BI-RADS Kategorie gibt. Die Überprüfung des Nutzens der Matrix war eine der Hauptzielsetzungen dieser Arbeit.

4.2 Diskussion der Ergebnisse

4.2.1 Histologisch-Pathologische Ergebnisse

Die von uns gefundenen histologischen Ergebnisse zeigen hinsichtlich der malignen Diagnosen ein eindeutiges Bild. Mit 89 Fällen liegt die Diagnose des DCIS deutlich an erster Stelle. Als zweit- und dritthäufigste Diagnosen finden sich die Mastopathie und die Adenosen Das IDC liegt mit 35 Fällen an vierter Stelle. Insgesamt wurde 138-mal die Diagnose eines malignen Befundes gestellt. Das DCIS und das IDC stellen gemeinsam über 90 % aller Malignome. Bei 102 der 203 Patientinnen zeigten die gewonnenen Histologien eine maligne Entität. Bei 36 Patientinnen lagen mehrere unterschiedliche Diagnosen gemeinsam vor. Hier ist mit 22 Fällen insbesondere das gemeinsame Vorliegen von DCIS und IDC zu nennen.

Das DCIS ist die häufigste karzinomatöse Veränderung der Brust, die durch das Mammographie Screening entdeckt wird. Dies begründet sich durch die typische Assoziation des DCIS mit Mikroverkalkungen [65, 111]. So sind nahezu 90 % der diagnostizierten DCIS zum Zeitpunkt der Diagnose durch Mammographie und Biopsie noch klinisch okkult. Die DCIS- Subtypen zeigen nach Exzision eine unterschiedlich hohe Rezidivgefahr [112]. Jedoch sind die unterschiedlichen Verkalkungsmuster nur moderat mit den verschiedenen pathologischen Subtypen des DCIS in Verbindung zu bringen [113] und somit von geringer prognostischer Wertigkeit.

Charakteristische und verdächtige Verkalkungen sind entscheidende mammographische Marker, die als Kriterium zur Biopsie herangezogen werden. Verkalkungen, insbesondere Mikroverkalkungen, sind die häufigsten mammographischen Veränderungen, die einer Biopsie zugeführt werden. Vor allem bei okkulten Karzinomen können Mikroverkalkungen den entscheidenden diagnostischen Hinweis geben. Ungefähr 25-43 % aller nicht palpablen Karzinome werden aufgrund von Mikroverkalkungen entdeckt [37, 114].

Eine Vielzahl an Arbeiten untersuchen in welcher Häufigkeit ein maligner Befund bei Mikroverkalkungen zu erwarten ist. Sie zeigen, dass in 21-47 % die Diagnose einer malignen Veränderung gestellt wird [39, 115-118].

Äthiologisch nimmt man an, dass das IDC sich aus Vorstufen entwickelt und somit aus dem DCIS, aus einer atypisch duktalem Hyperplasie (ADH) und aus duktalem Hyperplasie (DH) entstehen kann [17, 119]. In unserer Studie sind von den 35 Fällen eines IDC bei 22 bzw. 89% auch eine DCIS Komponente zu finden. Cady et.al fanden in ihrer Studie sogar bei 82 % aller IDC Fälle ein DCIS [119]. Stomper et al. fanden bei 32 % aller diagnostizierten DCIS eine IDC-Komponente [120].

Andererseits findet sich bei 26-38 % aller biopsierten malignen Mikroverkalkungen auch ein Fokus mit einem invasiven Karzinom [91, 112].

Die Gefahr der Invasion nimmt dabei mit dem pathologischen Malignitätsgrad des DCIS zu [121]. Diese Tatsache macht die genaue histologische Aufarbeitung zu einer obligaten therapie- und prognoseentscheidenden Bedingung, da die Entdeckung und Exzision des DCIS eine wichtige Strategie zur Prävention eines IDC ist [119].

In unseren histologischen Ergebnissen finden sich neun Fälle eines ILC. Ein LCIS wurde nicht gefunden. In zwei Fällen zeigte sich jedoch die Vorstufe eines LCIS in Form einer atypisch lobulären Hyperplasie (ALH).

Das LCIS und ILC ist nur sehr selten mit verdächtigen Mikrokalkarealen in Mammographien in Verbindung zu bringen [23, 122-125] und unsere Ergebnisse sind insofern nicht überraschend.

Da das ILC auch häufig keinen Herdbefund zeigt [123, 124, 126], ist die

mammographische und klinische Diagnose bzw. Früherkennung eine Herausforderung und bedingt dadurch seine Gefährlichkeit [122]. Daher werden die lobulären Karzinome selten in frühen Stadien entdeckt und häufig liegen bei Diagnose schon Lymphknotenmetastasen vor [122].

Die Patientinnen in dieser Studie wurden ausnahmslos im Vorfeld einer Biopsie zugeführt. Somit war vorab mindestens eine Kategorie 4 nach BI-RADS vergeben worden. Dies muss im Bezug auf den Aufbau der Studie als kritisch betrachtet werden. Denn somit ergibt sich eine Vorselektion und daraus resultierend eine überproportionale Anzahl an malignen Histologien. Darüber hinaus war diese Tatsache auch den befundenden Radiologen bekannt und es könnte angenommen werden, dass in Folge dessen eher eine höhere BI-RADS Kategorie vergeben wurde. Bei der Auswertung zeigt sich jedoch, dass eher das Gegenteil der Fall ist und insgesamt niedriger eingestuft wurde.

Insgesamt nehmen die malignen Histologien in unserer Studie 44 % aller gefundenen Diagnosen ein.

In anderen Studien bewegen sich die Zahlen maligner Histologien zwischen 21 und 47 % der Fälle bei unterschiedlich hohen Fallzahlen [106, 114, 118, 127, 128]. Im Vergleich dazu decken sich also unsere histologischen Ergebnisse mit denen anderer Untersuchungen.

Bei den benignen Diagnosen fallen 80 % auf fibrös- und fibrös-zystische Mastopathie, Adenosen, DH und apokrine Metaplasie. Diese Veränderungen sind nach unseren Ergebnissen die wichtigsten Differentialdiagnosen, die bei der Beurteilung von Mikroverkalkungen in Betracht gezogen werden müssen. Andere Studien kommen zu ähnlichen Ergebnissen. So sind die fibrös-zystische Mastopathie und die DH bei Yunus et al. [108] und de Lafontan [87] ebenfalls die häufigsten beschriebenen benignen Ursachen von Mikrokalk.

4.2.2 Statistische Ergebnisse

4.2.2.1 Sensitivitäten und Spezifitäten

Der Vergleich der Sensitivitäten und Spezifitäten zeigt signifikante Unterschiede zwischen den drei Methoden.

Die Sensitivität der modifizierten Matrix beträgt über 96 %. Dieses Ergebnis liegt damit mehr als 10 % über dem Ergebnis der subjektiven Einteilung und nahezu 16 % über dem der nicht modifizierten Matrix.

Auffallend sind besonders die ähnlichen Ergebnisse beim Vergleich der Untersucher untereinander. Alle vier Radiologen erzielten zwar kollektiv bei der Einteilung mit der Matrix die niedrigste Sensitivität, jedoch mit erheblichen Unterschieden untereinander. So liegt Untersucher 4 mit 93,14 % mehr als 23 % über dem Ergebnis von Untersucher 2. Diese Unterschiede zwischen den Readern zeigen sich in gleicher Weise auch bei den Sensitivitäten der anderen beiden Methoden, wenn auch nicht so ausgeprägt wie mit der ersten Methode, dem Matrixschema. Bei den Methoden 2, der subjektiven Auswertung, und 3, der modifizierten Matrix, sind die Ergebnisse näher beieinander. Untersucher 2 hat im Vergleich jeweils die niedrigsten Sensitivitäten. Man erkennt also durch alle Methoden hindurch eine Abstufung von Untersucher 4 mit den höchsten Sensitivitäten über Untersucher 3 und Untersucher 1 zu Untersucher 2, mit den niedrigsten Sensitivitäten. Bei den Ergebnissen der modifizierten Matrix steht zwar Untersucher 3 an erster Stelle mit der höchsten Sensitivität, jedoch ist der Unterschied zu Untersucher 4 sehr gering und beträgt nur wenige Prozentpunkte, sodass auch hier die Tendenz erhalten bleibt.

Eine der Fragen, die mit dieser Arbeit beantwortet werden sollte, ist ob die Matrix eine Hilfestellung für eher unerfahrene Befunder sein kann. Genauer bedeutet dies, dass man davon ausgeht, dass erfahrenere Radiologen durch mehrjährige und längere Erfahrung bessere Ergebnisse erzielen als Berufsanfänger. Dies müsste sich in den Ergebnissen bei der subjektiven

Beurteilung zeigen, da diese ja auf der Berufserfahrung aufbaut.

Nun haben Untersucher 2 und 1, wie in Kapitel 2 erläutert, mehr Erfahrung im Befunden von Mammographien als Untersucher 3 und 4.

Die weniger erfahreneren Untersucher zeigen jedoch nicht nur bei der subjektiven Einteilung höhere Sensitivitäten als die Erfahrenen sondern auch bei den anderen beiden Methoden.

Eine wahrscheinliche Begründung könnte sein, dass unerfahrenere Radiologen, eher dazu tendieren, unklare Fälle mit einer höheren BI-RADS Kategorie zu versehen. Dadurch erhöht sich die Wahrscheinlichkeit mehr tatsächlich Kranke, beziehungsweise richtig positive Fälle darunter zu haben. Diese Unsicherheit und auf ‚Nummer sicher gehen‘ - Taktik von Anfängern im Befunden von Mammographien ist gut nachzuvollziehen, da mit dem Einteilen in eine BI-RADS Kategorie auch immer eine Therapiestrategie verknüpft ist. Diese Problematik offenbart sich auch besonders bei den BI-RADS Kategorien 3 und 4. Hier verläuft die Grenze zwischen ‚follow-up‘ Untersuchung nach 6 Monaten und Biopsie, also zwischen Abwarten und Gewissheit haben.

Das würde aber bedeuten, dass sich in den Ergebnissen zeigen muss, das Unerfahrene tendenziell mehr Fälle in Kategorie 4 einteilen. Dass dem nicht so ist, zeigen die Ergebnisse der Auswertung der BI-RADS Kategorien. Darauf wird im Rahmen der Diskussion der Ergebnisse nach BI-RADS Kategorie weiter unten im Text wird darauf zurückgekommen.

Mit der subjektiven Einteilung haben alle vier Radiologen höhere Sensitivitäten als mit der Matrix. Jedoch haben alle vier Befunder noch einmal höhere Sensitivitäten mit der Einteilung nach der modifizierten Matrix. Die Einteilungshilfe mit der modifizierten Matrix übertrifft somit die Einteilung durch das subjektive Beurteilen. Dies gilt sowohl für erfahrene wie auch weniger erfahrene Radiologen. Die nicht modifizierte Matrix kann mit den niedrigsten Sensitivitäten jedoch den Erwartungen nicht standhalten.

Die Ergebnisse der Spezifitäten zeigen dagegen eine genau entgegengesetzte Verteilung beim Vergleich der Methoden. Die modifizierte Matrix hat hier mit 24% die niedrigste Spezifität und die nicht modifizierte Matrix mit 54% die

Höchste. Die subjektive Einteilung liegt auch hier dazwischen und kommt auf eine Spezifität von 51 %. Insgesamt fallen die Ergebnisse der Spezifität also deutlich schlechter aus, als die der Sensitivität.

Das begründet sich möglicherweise daraus, dass die beurteilenden Radiologen immer auch die Konsequenzen für den Patienten vor Augen haben.

Der Fehler der ‚Überdiagnostik‘ bei falsch- positiven Patienten hat meist weniger Konsequenzen als die Gefahr der Verschleppung einer Karzinomdiagnose eines falsch- negativen Patienten.

Ein bekanntes Problem ist die Tatsache, dass die Mammographie in der Detektion von Mikroverkalkungen zwar sehr sensitiv, jedoch die Spezifität nur gering ist [117]. Als Ergebnis finden sich eben die hohe Anzahl falsch- positiver Mammogramme und eine relativ niedrige Rate richtig positiver Biopsieraten [117].

Beim Vergleich der vier Untersucher untereinander fällt auf, dass auch bei den Ergebnissen der Spezifität die gleiche Reihenfolge der Untersucher untereinander wie bei der Sensitivität eingehalten wird. Untersucher 2 erreicht jeweils die höchsten Sensitivitäten. So sind die Ergebnisse von Untersucher 2 bei allen drei Methoden zwischen 20 und 30 % höher als bei Untersucher 4, der jeweils die niedrigsten Spezifitäten vorweist. Die Tendenz der Untersucher von höchster zu niedrigster Spezifität läuft also von Untersucher 2 über 1 und 3 zu Untersucher 4. Insgesamt sind die Ergebnisse weniger eng beieinander als bei den Ergebnissen der Sensitivität.

Der oben genannte Aspekt, dass Radiologen mit weniger Erfahrung unsicherer beim Befunden und Entscheiden sind, kann man auch mit den Ergebnissen der Spezifität vereinbaren. Untersucher 4 mit den niedrigsten Spezifitäten und den wenigsten Jahren an Erfahrung hat viel mehr falsch- positive Fälle als Untersucher 2.

Es fällt hierbei auf, dass die Ergebnisse der Untersucher in Bezug auf ihre Erfahrung spiegelbildlich verlaufen. Die Untersucher mit mehr Erfahrung zeigen bessere Spezifitäten und die mit weniger Erfahrung bessere Sensitivitäten.

Die Ergebnisse der Spezifität der Matrix liegen nur knapp 2 % über dem der

subjektiven Einteilung und somit auf vergleichbarem Niveau. Mit fast 30 % Differenz liegt die modifizierte Matrix an letzter Stelle der drei Methoden. Die Einteilung der Matrix ist dementsprechend keine Hilfe um bessere Spezifitäten zu erreichen.

Interessant sind die konträren Ergebnisse von modifizierter und nicht modifizierter Matrix. Während die modifizierte Matrix die signifikant höchste Sensitivität erreicht, fällt ihre Spezifität am niedrigsten aus. Die nicht modifizierte Matrix erreicht zwar knapp die höchste Spezifität, hat aber von allen Methoden die niedrigste Sensitivität. Dabei kann man dieses Ergebnis als Erklärung der Modifikation sehen, denn diese schiebt Fälle von Kategorie 3 zur 4 und bezeichnet somit mehr Mikroverkalkungen als Karzinom bzw. karzinomverdächtig.

4.2.2.2 Vergleich der Readings

Die Ergebnisse in Tabelle 38 und die Ergebnisse des Vergleich der Readings, zeigen sowohl eine Inter- wie auch eine Intrareadervariabilität.

Die Intrareadervariabilität ist wesentlich signifikanter als die Interreadervariabilität, die vernachlässigt werden kann, und zeigt sich vor allem in den Kategorien 3 und 4.

Es ist auffallend, dass mit der Matrix, trotz insgesamt leicht geringeren Unterschieden zwischen den Untersuchern, ebenfalls eine hohe Variabilität besteht. Auch die Interreadervariabilität zeigt sowohl bei der subjektiven Einteilung wie auch mit der Matrix ähnliche Abweichungen. Keine der Methoden zeigt hier Verbesserungen der Variabilität.

Am geringsten ausgeprägt ist die Inter- und Intrareadervariabilität bei der Einteilung in Kategorie 5. Bei Kategorie 2 sind sich 3 der 4 Untersucher mit 10-19 Fällen bei subjektiver bzw. 10-28 Fällen bei Einteilung mit der Matrix recht einig. Lediglich Untersucher 2 unterscheidet sich von den anderen Untersuchern, da er mit 89 beziehungsweise 72 Fällen beträchtlich mehr Fälle in diese Kategorie einstuft.

Auch in anderen Studien zeigen sich bei der Beschreibung der Morphologie von

Mikroverkalkungen und der radiologischen Klassifikation nach BI-RADS nur geringe oder mäßige Intra- und Interreader Übereinstimmungen [91, 97, 106].

Diese Ergebnisse und vor allem die geringe Übereinstimmung in den Kategorien 3 und 4 zeigen, dass Fälle mit Mikroverkalkungen, die weder hoch malignitätsverdächtig noch klar benigne sind, besonders schwierig zu kategorisieren sind [106].

Es fällt hierbei auf, dass Untersucher 4 im Gegensatz zu den anderen drei Readern überdurchschnittlich viele Fälle in Kategorie 4 eingeteilt hat. Mit der Einteilung durch die Matrix verringert sich die Zahl zwar um 60 Fälle, aber der deutliche Abstand zu den anderen Readern bleibt auch hier erhalten. Entsprechend der weiter oben genannten These bestätigt sich also, dass der Untersucher mit den geringsten Jahren an Erfahrung auch die meisten Kategorie 4 Fälle hat.

4.2.2.3 Übereinstimmungen

Betrachtet man die sowohl richtig- positiven wie auch die richtig- negativen Übereinstimmungen aller Untersucher gemeinsam, so schneidet die subjektive Einteilung, mit 69% im Gegensatz zur Methode mit der Matrix bzw. der modifizierten Matrix mit 67% und 61%, am besten ab. Dies impliziert, dass subjektive Einschätzung eines mammographischen Befundes, die ja auf die Erfahrung der befundenden Radiologen beruht, mehr als eine neutrale Matrix aussagt. Der Nutzen der Matrix muss eigentlich darin liegen die Schwächen einer subjektiven Einschätzung zu verringern. Durch die Interpretationshilfe der Matrix sollen mehr richtig- positive und richtig- negative Ergebnisse möglich sein.

Nun zeigen aber die Ergebnisse deutlich, dass die Matrix eben nicht diesen Anforderungen entspricht. Ganz im Gegenteil, sie ist weniger gut als die subjektive Einteilung und kann damit den an sie gestellten Erwartungen nicht erfüllen.

Die Befundung von Röntgenbildern, insbesondere von Mammographien, die sich durch eine Vielzahl unterschiedlicher und komplexer Befunde auszeichnet

erwartet unbestritten eine Menge Erfahrung von den Radiologen.

Es zeigte sich, dass die einzelnen Untersucher mit jeweils unterschiedlichen Methoden ihre besten Ergebnisse erzielten. So haben Untersucher 1 und Untersucher 4 mit der Matrix die höchsten Übereinstimmungen. Untersucher 2 und 3 dagegen mit ihrer subjektiven Einschätzung. Mit der modifizierten Matrix haben alle vier Untersucher die wenigsten Übereinstimmungen erzielt. Untersucher 4 hat die wenigsten Jahre an Erfahrung und Untersucher 2 die meisten. So liegt die Vermutung nahe, dass eher Unerfahrene von den Matrix profitieren und Erfahrene von ihrer eigenen subjektiven Einschätzung. Nun ist Untersucher 1 jedoch der Zweiterfahrene und erzielt mit der Matrix die meisten Übereinstimmungen. Untersucher 3 hat nach Untersucher 4 die wenigsten Jahre an Erfahrung und erzielt doch mit der subjektiven Einteilung die besten Ergebnisse. Insofern ist diese Vermutung wenig schlüssig.

Ursächlich zeigt sich die Tatsache, dass die vier Untersucher keine großen Erfahrungsunterschiede haben (siehe Kapitel 1) und daher mit Ihrer Befundungskompetenz zu nah beieinander liegen. Auch sind alle in der gleichen Abteilung und haben durch Erfahrungsaustausch voneinander gelernt, was zur Wissensgleichheit beigetragen hat. Pjinappel et. al. kritisiert Ähnliches in seiner Studie und vermutet, dass mit Radiologen anderer Abteilungen auch andere Ergebnisse erzielt worden wären [106].

Auffallend ist, dass Untersucher 4, auch wenn er mit der Matrix die meisten richtig- positiven und -negativen Ergebnisse hat, im Vergleich mit den anderen drei Readern bei allen drei Methoden am schlechtesten abschneidet. Kann man mit der Matrix also nur dann bessere Ergebnisse erzielen, wenn man auch schon mehr Erfahrung hat? Das würde dann aber auch bedeuten, dass nicht nur die Zuordnung zu den BI-RADS- Kategorien ein Problem ist, sondern auch schon die Bestimmung der Form und Verteilung. Die Erkennung selbst erfordert ja bereits ein gewisses Maß an Erfahrung. Die Komplexität der Erscheinungsformen von Mikroverkalkungen, deren Beschreibung und Erkennung bleibt also weiterhin das Hauptproblem bei der mammographischen Befundung.

4.2.2.4 Kontroverse BI-RADS Kategorie 3

In den Auswertungen unserer Ergebnisse haben wir die sicher und wahrscheinlich benignen BI-RADS- Kategorien zwei und drei zusammengefasst und als „benigne“ behandelt. Korrekterweise ist dies nur bedingt richtig, da in der BI-RADS- Kategorie drei zu einem gewissen Prozentsatz mit malignen Fällen gerechnet werden muss. Im BI-RADS Lexikon werden Werte zwischen 1-3 % angegeben und eine mammographische Kontrolle nach sechs Monaten empfohlen [70].

Wir finden dagegen je nach Methode Werte zwischen 8% und 13% bei Kategorie zwei und zwischen 19% und 32% bei Kategorie drei. Diese Werte unterscheiden sich erheblich von den im Lexikon und in der Literatur genannten Karzinomwahrscheinlichkeiten [70, 80, 129-131].

Diese Ergebnisse zeigen deutlich das zu Beginn dieses Kapitels erwähnte Problem, dass im BI-RADS Lexikon zu den Befunden keine Zuordnungsempfehlungen gemacht werden können, insbesondere dass hier ‚selektierter MK‘ biopsiert wurde. Damit ist gemeint, dass der Mikrokalk ja schon als verdächtig aufgefallen war und die Patientin gezielt zur Abklärung überwiesen wurde. Außerdem galt der MK in der Histologie einer Biopsie ja als Einschlusskriterium. Man kann also davon ausgehen, dass in dieser Studie eine selektierte Auswahl an Mikroverkalkungen vorliegt, die die Auswertung und Beurteilung der Ergebnisse zusätzlich erschwert. Durch diesen Selektionsbias entsteht ein systematischer Fehler, der die Beurteilung der Ergebnisse verzerrt. Mit der Matrix, also der nicht-subjektiven Beurteilung, ist zwar die Anzahl der malignen Fälle in den beiden benignen Kategorien teilweise niedriger als mit der subjektiven Einteilung, jedoch wäre es wünschenswert gewesen eine deutlichere Tendenz zu sehen.

In der Literatur wird die BI-RADS Kategorie 3 seit Beginn kontrovers diskutiert. Manche Autoren schlagen vor, auch die in Kategorie 3 eingeteilten Fälle zu biopsieren, da die klinische Aussagekraft durch eine hohe Intraindividuelle Variabilität nur begrenzt ist [106, 129]. Besonders trifft dies zu, wenn die Kategorie

aufgrund von gefundenen Mikroverkalkungen vergeben wurde.

Weiterhin werden auch oft, im Vergleich mit Voraufnahmen, stabile Mikroverkalkungen in Follow-up-Untersuchungen für wahrscheinlich benigne kategorisiert. Lev-Toaff et al. fand in seiner Studie jedoch in dieser Gruppe ein Malignitätsrate von 24,8% [132].

Andere Autoren betonen aber genau den Vorteil dieser Kategorie dadurch, dass bei richtiger Anwendung damit unnötige Biopsien vermieden werden können [131, 133, 134]. Sie unterstreichen die Notwendigkeit der Definition von standardisierten und allgemein akzeptierten Kriterien für die Einteilung in die wahrscheinlich benigne BI-RADS Kategorie drei, die auch bisher nicht eindeutig gelungen ist [131].

Zusammenfassend lässt sich sagen, dass auch in unseren Ergebnissen eine hohe Intrareadervariabilität der BI-RADS Kategorien drei und vier, vorliegt.

Die Schwierigkeit einer klaren Definition von Kriterien, wie in unserer Studie z.B. bei Mikroverkalkungen, bleibt auch im speziellen Fall der umstrittenen Kategorie 3 bestehen und kann trotz Matrix nicht verbessert werden.

4.2.2.5 Ergebnisse nach BI-RADS- Kategorie

In Kapitel 3.1.7 werden die Ergebnisse je nach BI-RADS Kategorie aufgeführt und die drei Methoden diesbezüglich miteinander verglichen.

Zahlreiche Studien belegen, dass die BI-RADS Kategorien zur Prognose der Malignität herangezogen werden können [78, 79, 91, 95, 98, 100, 135]. Die Karzinomrate und die Erkrankungswahrscheinlichkeit beziehungsweise der positive Prädiktwert (=PPV) nehmen von Kategorie 2 zu 5 kontinuierlich zu [78, 79, 91, 95, 98, 100, 135].

Auch bei unseren Ergebnissen zeigt sich sowohl bei der subjektiven Einteilung wie auch bei der Einteilung mit der Matrix eine Zunahme der malignen Fälle von BI-RADS Kategorie zwei zu fünf.

Die Kategorien zwei und drei, in die sicher bzw. wahrscheinlich benigne Befunde klassifiziert werden, zeigen jedoch hohe Anteile an malignen Fällen.

Bei der subjektiven Einteilung sind 13% maligne Fälle als sicher benigne und 24% als wahrscheinlich benigne klassifiziert worden. Mit der Matrix wurden sogar fast ein Drittel, der in BI-RADS Kategorie 3 eingeordneten Fälle, als maligne beurteilt.

Wären diese Fälle nicht einer Biopsie zugeführt worden, hätte dies eventuell gravierende Folgen für die betroffenen Patientinnen gehabt.

Die im Vergleich den tatsächlich Gegebenheiten am nächsten kommenden Ergebnisse in den beiden ‚benignen‘ Kategorien wird von der modifizierten Matrix erreicht. Hier findet man in Kategorie zwei 8% und in Kategorie drei 19% der malignen Fället.

Bei den wahrscheinlich malignen bzw. malignom- typischen Kategorien vier und fünf zeigt wieder herum die subjektive Einteilung mit 53% in Kategorie 4 und 86% maligne Fälle in Kategorie 5 die am nächsten kommenden Ergebnisse entsprechend den vom ACR veranschlagten Wahrscheinlichkeiten.

Die Studie von Pjinappel et al. [106] zeigt ähnlich hohe Malignitätsraten in BI-RADS Kategorie 3 wie in unseren Ergebnissen. In seiner Studie untersucht er Mammographien mit nicht palpablen Befunden, hauptsächlich Mikroverkalkungen. In einer ersten Sitzung fand er 16% maligne Fälle, in der zweiten Sitzung sogar 18%.

Die meisten anderen Studien zeigen jedoch übereinstimmend geringere Werte. Auch Obenauer et al. untersuchte Mikroverkalkungen, die als BI-RADS Kategorie 3 klassifiziert wurden und fand einen PPV von etwa 4% [96].

Varas et al. [130] und Sickles et al. [80] verkündeten in ihren Arbeiten, dass sich von den gefundenen wahrscheinlich benignen Befunden 2% oder weniger als tatsächlich maligne herausgestellt haben. [130]

Eine von Varas et al. im Jahre 2002 durchgeführte Folgestudie zur Überprüfung seiner zehn Jahre zuvor gefundenen Ergebnisse zeigen sogar noch niedrigere Werte. Der Positive Prädiktivwert in Kategorie drei liegt bei 0,4% [136]. Mendez et al. untersuchte die Rate an falsch-negativen Befunden in BI-RADS 3 Läsionen. Das Ergebnis zeigte, dass 4,5 % der Mammographien als falsch-negativ befundet wurden. Darunter waren die höchsten Karzinomraten bei Patienten mit Mikroverkalkungen [137].

Die von uns gefundenen Werte stimmen aber insgesamt wenig mit den vom ACR im Lexikon veranschlagten oder von anderen Autoren veröffentlichten Werten überein. Besonders deutlich wird dies in Kategorie 2 und 3. Bei allen drei Methoden ist mit einer deutlich höheren Karzinomwahrscheinlichkeit zu rechnen.

In unserer Studie werden die Mammographien nur anhand von Mikroverkalkungen eingeordnet. Die meisten der oben genannten Studien generieren Ihre Klassifizierung aus anderen Kriterien und ziehen zur Klassifizierung auch andere mammographische Befunde oder Untersuchungsmethoden wie die Sonographie mit ein, was die Vergleichbarkeit der Ergebnisse erheblich einschränkt.

Gleichzeitig zeigt dies aber auch wie wichtig Definitionskriterien zur Klassifizierung sind und dass mit unterschiedlichen Kriterien auch unterschiedliche Ergebnisse zu erwarten sind [137].

Bilder mit Opazitäten wurden aus der Studie ausgeschlossen um zu gewährleisten, dass die Ergebnisse nur aufgrund der Einteilung von Verkalkungen zustande kommen. Nicht ausgeschlossen ist es jedoch, dass der die Patientin zuweisende Kollege etwas Verdächtiges in der sonographischen Untersuchung gesehen hat, was vor allem bei BI-RADS 2 Verkalkungen anzunehmen ist.

Daraus lässt sich schließen, dass man zur korrekten Klassifizierung von Mammographien alle diagnostischen Befunde mit einbeziehen sollte oder aber, dass der Versuch einer Einteilung nur durch Mikroverkalkungen doch komplexer ist als angenommen.

Ein interessantes Ergebnis findet sich auch in der oben genannte Studie von Varas et al. [136]. So liegt der PPV zwar bei 0,4%, aber bei Patienten, die aufgrund einer Befundprogression biopsiert wurden, zeigte sich eine Malignitätsrate von 14%.

Wie in Kapitel 2 beschrieben sind die von uns in die Studie eingeschlossenen Patienten ein sehr heterogenes Kollektiv. Unter den Patienten fanden sich solche mit einer Screening Mammographie bei unauffälligem Untersuchungsbefund, Patienten mit verdächtigem Tastbefund oder Patienten

mit Follow-up- Mammographie nach BI-RADS 3 Läsion. Bei letzteren kann eine solche Befundprogression vorgelegen haben, wodurch sich höhere Karzinomwahrscheinlichkeiten erklären lassen.

Insofern scheint ein Vergleich mit anderen Studien wenig sinnvoll, da die meisten anderen Studien mit homogeneren Patientenkollektiven gearbeitet haben.

4.2.2.6 Verteilung der Malignitätswahrscheinlichkeit über die Matrix

Die Verteilung der Malignitätswahrscheinlichkeit über die Matrix zeigt tendenziell eine Zunahme derselben von diffuser zu linear- duktaler Verteilung und von typisch benigner zu typisch maligner Verkalkungsform. Dies ist auch im BI-RADS Lexikon so beschrieben.

Betrachtet man nur die Verkalkungsformen alleine, so zeigen die als typisch benigne eingestuftten Verkalkungen mit Malignitätswahrscheinlichkeiten von 4bis 23% die niedrigsten Werte. Die typisch malignen Verkalkungen zeigen mit 37%- 88% die höchsten Werte. Mit 14%-70% liegen die intermediären Verkalkungen dazwischen. Diese Verteilung entspricht also den erwartenden Ergebnissen beziehungsweise den am häufigsten zugrunde liegenden Ursachen der entsprechenden Verkalkungen.

Die von uns als typisch maligne eingestufte Verkalkungen sind wie in Kapitel eins beschrieben, pleo- und polymorphe, verzweigende, y- förmige, irreguläre, heterogen- granuläre, feine Verkalkungen. Mehrere Studien beschäftigen sich mit der Frage nach der diagnostischen Wertigkeit von Mikroverkalkungen als mammographischer Marker und bestätigen die oben genannten Formen von Mikroverkalkungen als Hinweise für eine maligne Histologie.

Nach der Studie von Yunus et.al sind ebenfalls verästelnde, lineare, irreguläre und y- förmige Verkalkungsformen spezifische und statistisch signifikante Kriterien von malignen Mikroverkalkungen [108].

In mehreren Studien wurde dieser Zusammenhang bestätigt [87, 91, 110, 138]. Le Gal et al. [139], der schon 1984 den Versuch einer Klassifizierung und

Einteilung von Mikroverkalkungen in Bezug auf deren diagnostische Wertigkeit vornahm, unterschied 5 Gruppen:

Typ 1: ringförmige Verkalkungen. Typ 2: regelmäßige, punktförmige Verkalkungen. Typ 3: feine, zarte Verkalkungen deren Form schlecht erkennbar ist. Typ 4: unregelmäßige punkt- und salzförmige Verkalkungen. Typ 5: wurmförmige, y-förmige Verkalkungen. Die Malignität nimmt von Typ 1 bis Typ 5 zu. Sie beträgt bei Typ 1 0% und bei Typ 5 100% .

Dieser starke Zusammenhang zwischen Le Gals Typ 5 und Malignität wurde in Folgestudien reevaluiert und bestätigt [87, 108]

Yunus et al. [108] fand auch einen Zusammenhang mit malignem Geschehen und der zunehmenden Anzahl von Mikroverkalkungen, insbesondere von mehr als 20 Verkalkungen pro cm².

Auch Le Gal et al. [139] und andere Studien [39, 87, 138] fanden einen positiven Zusammenhang mit der steigenden Anzahl von Verkalkungen und der malignen Histologie der Brustveränderung.

Dieser Aspekt wird in unserer Klassifikation von Mikroverkalkungen mit dem Matrixschema nicht beachtet und stellt daher ein Kritikpunkt dar. Bei der subjektiven Beurteilung jedoch ist er mit Sicherheit ein Faktor der entscheidend zur Einteilung beigetragen hat.

Bei den typisch benignen Verkalkungen fällt die rund/ovale Verteilung mit 15% aus dem im ersten Absatz genannten Muster. Zusammen mit der linear-duktalem Verteilung hat sie die höchste Malignitätswahrscheinlichkeit. Auffallend ist hierbei der große Unterschied zu diffusem, regionalem oder segmentalem Verteilungsmuster. Man könnte hierbei vermuten, dass eher Gruppen mit vielen Verkalkungen auf kleinem Raum in die rund/ovale Verteilung eingeteilt wurden und dass gemäß dem oben genannten dadurch die höhere Malignitätswahrscheinlichkeit zustande kommt. Um dies zu bestätigen, wäre jedoch eine erneute Evaluation notwendig.

Die linear- duktalem Verteilung ist auch bei typisch benigne aussehendem Einzelkalk in gut einem Viertel der Fälle malignomverdächtig. Wie man aus Tabelle 39 aber auch entnehmen kann sind die Fallzahlen mit insgesamt 13 Fällen niedrig, so dass eine genauere Verifikation erst mit höheren Fallzahlen

gelingt. Nichts desto Trotz sollte Mikrokalk mit linear- duktaler Verteilung auch im Hinblick auf intermediären oder typisch malignem Mikrokalk entsprechende Aufmerksamkeit beziehungsweise Sorgfalt bei der Prüfung zuteil werden und eventuell auch bei unauffälligem Mikrokalk einer Biopsie zugeführt werden.

So zeigen die linear- duktales Verkalkungen bei allen drei Gruppen jeweils die höchsten Malignitätswahrscheinlichkeiten und erreichen bei typisch-malignem Mikrokalk 88%.

Das diffuse Verkalkungsmuster, das eher auf eine benigne Pathologie schließen lässt [108] erreicht bei typischer maligner Morphologie des Einzelkalkes mit 91% den Höchstwert der Malignitätswahrscheinlichkeiten. Aber auch hier ist dies das Ergebnis einer geringen Fallzahl und stellt dadurch die Reproduzierbarkeit in Frage. Auch waren diese Fälle auf eher untypische, sehr ausgedehnte DCIS zurückzuführen. Zusammenhängen könnte dies mit der Tatsache, dass es sich beim Universitätsklinikum Tübingen um ein Referenzzentrum handelt, d.h. dass dort eher komplexe Fälle überwiesen und behandelt werden.

Yunus et al. [108] fand in seiner Studie bei 28,57% der Patienten mit diffuser Mikrokalkverteilung eine maligne Histologie. Auch er kommt zu dem Ergebnis, dass die Form der Einzelverkalkungen der wichtigste Parameter ist. Wenn diese ein irreguläres, heterogenes Aussehen haben besteht trotz diffuser Verteilung der Verdacht auf ein malignes Geschehen, dass durch Biopsie verifiziert werden sollte [140].

Nach Liberman et al. [91] sind lineare und pleomorphe Mikroverkalkungen hoch-spezifisch wenn nicht sogar pathognomonisch für DCIS.

Zusammenfassend lässt sich sagen, dass die Malignitätswahrscheinlichkeit von diffuser zu linear- duktaler Verteilung ansteigt und dass das Verkalkungsmuster umso verdächtiger wird, je verdächtiger und suspekter die Form des Mikrokalkes ist.

Vergleicht man unsere gefundenen Malignitätswahrscheinlichkeiten, mit denen, die sich nach der Zuordnung zu den BI-RADS Klassen in der Matrix ergeben so fällt auf, dass mit durchgehend höheren Malignitätswahrscheinlichkeiten zu rechnen ist als die in der Matrix zugeteilten. Besonders fällt dies bei den

typisch-benignen Verkalkungen auf. Nach unseren Ergebnissen hätte hier bei rund/ovaler und linearer Verteilung sogar eine BI-RADS Kategorie vier vergeben werden sollen. Die Matrix ordnet hier eine Kategorie zwei bzw. drei zu. Eine BI-RADS Kategorie zwei wäre nach unseren Ergebnissen bei keiner Zuordnung vergeben worden.

Die Modifikation der Matrix bei den intermediären Verkalkungen trifft die von uns ermittelten Malignitätswahrscheinlichkeiten besser als die nichtmodifizierte Form.

Die BI-RADS Kategorie 4 wurde in der neuesten Auflage des Lexikons von 2003 in drei Unterkategorien zerlegt: 4a, 4b und 4c. Die Malignitätswahrscheinlichkeit hat daher eine große Spannweite von 4%- 90%. Geht man von diesen Werten aus, so kann man die von uns ermittelten Ergebnisse mit den von der Matrix zugeteilten Vierer-Kategorien durchaus vereinbaren.

Auch die von der Matrix zugeordneten fünfer- Kategorien bei typisch-malignen Verkalkungen mit segmentaler und linearer Verteilung decken sich nahezu mit den von uns gefundenen Ergebnissen.

In der Literatur findet man, dass einige der Verteilungsmuster, wie zum Beispiel die Regionale im Gegensatz zur Segmentalen, keine wirklich signifikanten statistischen Unterschiede zeigen. Unsere Ergebnisse zeigen bei der segmentalen Verteilung jedoch einen höheren Anteil an malignen Fällen wie bei der regionalen Verteilung. Dies gilt für den typisch benignen, den intermediären wie auch den typisch malignen Kalk und zeigt dadurch die Unabhängigkeit von der Morphologie.

4.3 Abschließende Beurteilung und Ausblick

Die widersprüchlichen Ergebnisse unserer Studie zeigen, dass die Matrix in ihrer jetzigen Form noch nicht den von ihr erhofften Nutzen zeigt. Die Probleme beim Befunden von Mammographien mit Mikroverkalkungen, wie niedrige Sensitivitäten und Spezifitäten oder hohe Variabilitäten zwischen verschiedenen Befundern, konnten in unserer Studie durch die Matrix nicht wesentlich

reduziert oder verbessert werden.

Eine Hilfe für noch unerfahrene Radiologen konnte unter den Bedingungen in unserer Studie ebenfalls für die Matrix nicht belegt werden.

Jedoch zeigen einzelne Ergebnisse, wie die hohe Sensitivität der modifizierten Matrix und der hohe Anteil maligner Fälle in den ‚benignen‘ BI-RADS Kategorien, dass ein Überdenken und eine Überarbeitung der zugeordneten Kategorien beziehungsweise weitere Evaluationen dazu notwendig und sinnvoll wären.

Was sich in unseren Ergebnissen vor allem deutlich widerspiegelt ist, dass das Erkennen und Beschreiben der Mikroverkalkungen selbst, eines der Hauptprobleme ist. Das Beschreiben und die klare Definition von Verkalkungen ist der erste Schritt beim Befunden der Mammographien. Diese Problematik bringt auch die Matrix mit sich.

Seit nunmehr zwei Jahrzehnten versucht man mit Hilfe des Computers diese Schwierigkeiten beim Erkennen und Einteilen von Mikroverkalkungen zu unterstützen und zu vereinfachen. Der Begriff ‚computer-aided-detection‘ (=CAD) vereint eine Vielzahl unterschiedlicher Methoden und Systeme, die durch eine Computer unterstützte Technik die Diagnosefindung und Befundung von Mammographien erleichtern soll [141-143].

Neben Systemen für die Detektion von Verdichtungen sind ein Großteil der CAD- Systeme speziell für die Identifikation von Mikroverkalkungen entwickelt worden [1, 144-150]. Diese versuchen eine automatisierte Klassifikation von Mikroverkalkungen, in benigne und maligne Verkalkungen mit Hilfe verschiedener Methoden, wie zum Beispiel künstlich neuronale Netzwerke oder mathematische Systeme und Algorithmen zu erstellen.

Zahlreiche Studien belegen, dass durch Hilfe dieser Systeme eine höhere Spezifität und Sensitivität wie auch eine Verringerung der Interreadervariabilität erreicht werden kann [145, 151-154]. Jiang et. al entwickelten ein CAD-System für die Klassifizierung von Mikroverkalkungen [155]. Mit Hilfe dieses Systems erreichten die Radiologen eine signifikante Verbesserung der Klassifikation und eine Reduktion der Biopsieraten für benigne Läsionen.

Auch bei der Zuordnung zu den BI-RADS Kategorien zeigen Computerunterstützte Analysen bessere Ergebnisse in Form höherer PPV [156-158]. Besonders in den Kategorien drei und vier, in denen die Diagnosefindung besonders schwierig ist bzw. die Variabilität besonders hoch ist, kann die Interpretation mit Hilfe eines CAD-Systems die Leistung des Radiologen verbessern [156, 159].

Obwohl mit den CAD Systemen ein viel versprechender Weg beschritten wird, wird man auch hier noch mit verschiedenen Problemen konfrontiert [1, 109, 160-162]. Es bleiben somit Herausforderungen und Ziele für weitere Forschungen und Verbesserungen der CAD-Systeme bestehen.

Am besten überzeugen Kombinationen von verschiedenen Methoden. Gute Ergebnisse erzielen die CAD-Systeme, wenn sie nicht alleine sondern als Art zweite Befunder wie beim ‚double reading‘ eingesetzt werden [1, 53, 151, 154]. Die Diagnosefindung des Radiologen wird dabei unterstützt und erleichtert, aber nicht durch einen Computer ersetzt. Trotz neuer Techniken sollte die Befundung von Mammographien auch weiterhin nicht allein einem Computer überlassen werden.

Die vielfältige Problematik der Befundung von Mikroverkalkungen in Mammographien bleibt also auch trotz Entwicklung von CAD-Systemen in Zukunft noch ein Gebiet auf dem viele Fragen offen sind, auf dem es viel zu erarbeiten gibt und auf dem spannende Ergebnisse erwartet werden dürfen.

5. Zusammenfassung

Mikroverkalkungen spielen bei der Früherkennung von Brustkrebs eine sehr wichtige Rolle. Die Einteilung von Mikrokalk und ihre Zuordnung zu den BI-RADS Kategorien bereiten jedoch Schwierigkeiten, da nur wenige Daten dazu vorliegen und selbst im Lexikon des ACR keine Empfehlungen dazu gemacht werden. Diese Arbeit beschäftigt sich mit einer neuen matrix- ähnlichen Einteilung von Mikroverkalkungen. Diese gibt je nach Form und Verteilung der Verkalkungen eine Empfehlung für die Zuordnung zu einer BI-RADS- Kategorie. Die Zielsetzung dieser Arbeit ist die Überprüfung der Matrix auf ihren Nutzen und der Validität ihrer Zuordnungen hin.

Die digitalen Mammographiebilder von 203 Patientinnen mit Mikroverkalkungen wurden von 4 Radiologen zweimal im Abstand von mindestens vier Wochen befundet und eine BI-RADS Kategorie vergeben. Die Befundung wurde zum einen mit Hilfe der Einteilung durch die Matrix und eine Modifikation der Matrix und zum anderen durch die subjektive Einschätzung vorgenommen. Bei Einteilung mit der Matrix muss der Mikrokalk nach Morphologie in typisch benigne, intermediär oder typisch maligne und nach Verteilung in diffus, rund, oval, regional, segmental oder linear eingeordnet werden. Die Beschreibung der Verkalkungen und die subjektive Kategorievergabe richteten sich an die im BI-RADS Lexikon beschriebenen Vorgaben. Die modifizierte Matrix ordnet bei intermediären Verkalkungen andere BI-RADS Kategorien als die Ausgangsmatrix zu. Bei allen Patienten war das suspekte Mikrokalkareal im Vorfeld durch Biopsie und histologische Aufarbeitung diagnostisch untersucht worden. Eine Übereinstimmung mit der histologischen Diagnose galt bei benigner Histologie und BI-RADS Kategorie 2 und 3 beziehungsweise maligner Histologie und BI-RADS Kategorie 4 und 5. Die erhobenen Daten wurden statistisch ausgewertet.

Die häufigsten histologischen Diagnosen sind das Duktale Carcinoma in situ

und das intraduktale Karzinom mit zusammen 124 Fällen. Die häufigsten benignen Diagnosen sind die Mastopathie und die Adenosen mit insgesamt 100 Fällen.

Sowohl die subjektive Kategorievergabe wie auch die Einteilung mit der Matrix zeigen hohe Inter- und Intrareadervariabilitäten.

Die Ergebnisse der Sensitivität und Spezifität zeigen für die Matrix, die modifizierte Matrix und die subjektive Einteilung deutliche Unterschiede. Für die Matrix sind es 81% und 54%, für die modifizierte Matrix 97% und 24%, und für die subjektive Einteilung 86% und 52 %.

Die meisten Übereinstimmungen insgesamt, also sowohl die richtig positiven wie auch die richtig negativen Fälle werden mit der Einteilung nach subjektiver Einschätzung erreicht. Die Matrix selbst zeigt entsprechend ihren Vorgaben eine Zunahme der Malignitätswahrscheinlichkeit von diffuser zu linearer und von typisch benignem zu typisch malignem Mikrokalk. Die BI-RADS Kategorien zwei und drei zeigen sowohl mit der Matrix wie auch mit der subjektiven Einschätzung jedoch eine signifikant hohe Anzahl an malignen Fällen.

Die Ergebnisse dieser Arbeit zeigen, dass beim Befunden von Mikroverkalkungen in Mammographien, die Hauptschwierigkeit im Beschreiben und Erkennen der komplexen Mikroverkalkungen liegt. Darin sind die niedrigen Übereinstimmungen auch bei Einteilung mit der Matrix zwischen den Untersuchern untereinander und zwischen den beiden Readings begründet. Die Zuordnungen der BI-RADS Kategorien in der Matrix sollten überdacht und reevaluiert werden. Die Matrix zeigt zwar eine hohe Spezifität und die modifizierte Matrix eine hohe Sensitivität, jedoch unterliegen beide in ihrer jeweiligen Gesamtheit der subjektiven Einteilung.

6. Anhang

6.1 Tabellenverzeichnis

Tab. 1: Spezielle Tumore/Mammographie	9
Tab. 2: Gegenüberstellung ‚maligner - ‚benigner‘ Verkalkungen in der Mammographie.....	12
Tab. 3: Übersicht BI-RADS-System	14
Tab. 4: Matrixsystem	28
Tab. 5: Histologische Diagnosen und Fallanzahlen	32
Tab. 6: Gegenüberstellung Anzahl Diagnosen-Anzahl Patienten.....	34
Tab. 7: Sensitivität und falsch negative Befunde der Methode 1	35
Tab. 8: Signifikanztests der Sensitivitätsberechnung bei der 1.Methode	37
Tab. 9: Gesamtsensitivität und Konfidenzintervalle der 1.Methode.....	37
Tab. 10: Spezifität und falsch positive Befunde der Methode 1	38
Tab. 11: Signifikanztests der Spezifitätsberechnung der 1.Methode.....	39
Tab. 12: Gesamtspezifität und Konfidenzintervalle der 1.Methode	40
Tab. 13: Sensitivität und falsch negative Befunde der Methode 2.....	41
Tab. 14: Signifikanztests der Sensitivitätsberechnung bei der 2. Methode	42
Tab. 15: Gesamtsensitivität und Konfidenzintervalle der 1.Methode	42
Tab. 16: Spezifität und falsch positive Befunde der Method 2.....	44
Tab. 17: Signifikanztests der Spezifitätsberechnung der 2. Methode.....	44
Tab. 18: Gesamtspezifität und Konfidenzintervalle der 2.Methode	45
Tab. 19: Sensitivität und falsch negative Befunde der Methode 3.....	46
Tab. 20: Signifikanztests der Sensitivitätsberechnung bei der 2. Methode	47
Tab. 21: Gesamtsensitivität und Konfidenzintervalle der 3.Methode.....	48
Tab. 22: Spezifität und falsch positive Befunde der Methode 2.....	48
Tab. 23: Signifikanztests der Spezifitätsberechnung der 3. Methode.....	49
Tab. 24: Gesamtspezifität und Konfidenzintervalle der 3.Methode	50
Tab. 25: Signifikanztests der Gesamtsensitivitätsberechnung	52
Tab. 26: Signifikanztests der Gesamtspezifitätsberechnung	53
Tab. 27: Gesamtübereinstimmungen und Konfidenzintervalle der 1.Methode.....	59
Tab. 28: Gesamtübereinstimmungen der Methode 1 je Reader.....	60
Tab. 29: Signifikanztests der Berechnung der Gesamtübereinstimmungen der 1. Methode	60
Tab. 30: Gesamtübereinstimmungen und Konfidenzintervalle der 2.Methode.....	61

Anhang

Tab. 31: Gesamtübereinstimmungen der Methode 2 je Reader	62
Tab. 32: Signifikanztests der Berechnung der Gesamtübereinstimmungen der 2. Methode	63
Tab. 33: Gesamtübereinstimmungen und Konfidenzintervalle der 3.Methode	64
Tab. 34: Gesamtübereinstimmungen der Methode 3 je Reader	65
Tab. 35: Signifikanztests der Berechnung der Gesamtübereinstimmungen der 3. Methode ...	65
Tab. 36: Einteilung in die BI-RADS Kategorien durch die Matrix und die subjektive Einteilung.	67
Tab. 37: Einteilung in die BI-RADS Kategorien durch die modifizierte Matrix und die subjektive Einteilung.....	67
Tab. 38: Einteilungen in die BI-RADS Kategorien für alle Reader getrennt.	69
Tab. 39: Verteilung der Malignitätswahrscheinlichkeit über die Matrix	70
Tab. 40: Gegenüberstellung der Malignitätswahrscheinlichkeiten über die Matrix.....	71

6.2 Abbildungsverzeichnis

Abb. 1: Die zehn führenden Krebsarten	1
Abb. 2: T-Stadienabhängige Überlebensraten des Mammakarzinoms	3
Abb. 3: Anteil maligner Diagnosen	33
Abb. 4: Anteil benigner Diagnosen.....	33
Abb. 5: Sensitivität Methode 1.....	35
Abb. 6: Gesamtsensitivität der 1. Methode	37
Abb. 7: Spezifität Methode	38
Abb. 8: Gesamtspezifität der 1. Methode	39
Abb. 9: Sensitivität Methode 2.....	40
Abb. 10: Gesamtsensitivität der 2. Methode	42
Abb. 11: Spezifität Methode 2.	43
Abb. 12: Gesamtspezifität der 2. Methode	45
Abb. 13: Sensitivität Methode 3. Angaben	46
Abb. 14: Gesamtsensitivität der 3. Methode	47
Abb. 15: Spezifität Methode 3.	48
Abb. 16: Gesamtspezifität der 3. Methode.....	50
Abb. 17: Vergleich der Gesamtsensitivitäten.	51
Abb. 18: Vergleich des Gesamtspezifitäten.	52
Abb. 19: Vergleich des 1. und 2. Readings bei Reader 1	54
Abb. 20: Vergleich des 1. und 2. Readings bei Reader 2	55
Abb. 21: Vergleich des 1. und 2. Readings bei Reader 3.....	56
Abb. 22: Vergleich des 1. und 2. Readings bei Reader 4	57
Abb. 23: Gesamtübereinstimmungen der 1. Methode.	58
Abb. 24: Gesamtübereinstimmungen je Reader Methode 1	59
Abb. 25: Gesamtübereinstimmungen der 2. Methode	61
Abb. 26: Gesamtübereinstimmungen je Reader Methode 2.....	62
Abb. 27: Gesamtübereinstimmungen der 3. Methode.	63
Abb. 28: Gesamtübereinstimmungen je Reader Methode 3.....	64
Abb. 29: Beispielmammographie mit typisch segmental gruppiertem Mikrokalk.....	72
Abb. 30: Beispielmammographie mit im Fettgewebe auskristallisiertem Mikrokalk.....	72
Abb. 31: Beispielmammographie mit fibrös-zystischer Mastopathie.....	73
Abb. 32: Beispielmammographie mit DCIS, IDC und Fettgewebsnekrosen.....	73
Abb. 33: Beispielmammographie mit typisch malignem Mikrokalk.	74

Anhang

Abb. 34: Beispielmammographie mit linearem, pleomorpher Mikrokalk.....	75
Abb. 35: Beispielmammographie mit fibröser Mastopathie.....	76
Abb. 36: Beispielmammographie mit poly- und pleomorphem Mikrokalk.....	76
Abb.37: Beispielmammographie mit heterogenem, granulärem Mikrokalk.....	77

6.3 Abkürzungen

Abb.	Abbildung
ACR	American College of Radiology
ADH	Atypisch Duktale Hyperplasie
ALH	Atypisch Lobuläre Hyperplasie
BET	Brusterhaltende Therapie
BI-RADS	Breast Imaging-Reporting and Data System
CAD	Computer Aided Detection
CC	Cranio-Caudal
CI	Konfidenzintervall
DCIS	Duktales 'Carcinoma In Situ'
DIN	Duktales Intraepitheliale Neoplasie
DH	Duktales Hyperplasie
Et. al	et alii, und andere
FFDM	Full-Field-Digital-Mammographie
FSM	Film-Screen-Mammographie
IDC	Invasives Duktales Karzinom
ILC	Invasives Lobuläres Karzinom
LCIS	Lobuläres 'Carcinoma In Situ'
LIN	Lobuläre Intraepitheliale Neoplasie
LK	Lymphknoten
MK	Mikrokalk
ML	Medio-lateral
MLO	Medio-lateral-oblique
MO	Molybdän
MRT	Magnet-Resonanz-Tomographie
PPV	Positive Predictive Value
RHO	Rhodium
Tab.	Tabelle
TNM	Tumor-Nodus-Metastasen
VB	Vakuumbiopsie
WHO	World Health Organisation

6.4 Abstract

Background

Although much progress has been made in early detection, diagnosis and therapy of breast cancer in the past few years, it is still one of the leading causes of cancer death among women. Mammography is the mainstay of early diagnosis of breast cancer. Microcalcifications in mammograms play a very important role in the detecting of breast cancer, especially early stage; however they are not breast cancer specific. Studies report, that mammography show microcalcification in up to nearly 50% of cases of breast cancer. The BI-RADS Lexicon defines parameters to classify the mammographic features into standardized diagnostic categories. Yet, the classification of microcalcification raises difficulties, particularly because there are no recommendations in the Lexicon on how to link the descriptive findings to the BI-RADS assessment categories.

Purpose

The purpose of our study is to evaluate a new kind of classification scheme of microcalcification. This newly developed matrix-like scheme combines the morphology and distribution parameters of microcalcification and gives for each combination an assignment for the BI-RADS assessment categories.

Material and Methods

The digital mammography images of 203 patients with microcalcification were twice interpreted by 4 Readers at intervals of at least four weeks. They classified the images on the one hand according the classification scheme and on the other hand according to the BI-RADS categories.

The classification process with the 'matrix' scheme divides the morphology into three groups: typically benign, intermediate and typically malignant calcifications. The distribution is divided into five groups: diffuse, round or oval, regional, segmental and linear.

All patients underwent vacuum core biopsy or open surgery. The histological findings were compared to the two radiological classifications. BI-RADS categories 2 and 3 were considered as representing benign cases, whereas BI-

RADS categories 4 and 5 were counted as malignant cases. The significance of the correlation between the Bi-RADS classification and the classification with the matrix scheme and the diagnostic indices (sensitivity, specificity, positive predictive value) were calculated using standard statistical methods.

Results

The most frequent histological results were observed in 124 malignant cases in form of DCIS and invasive ductal cancer. The most frequent benign diagnoses were fibrocystic changes and adenosis.

The classification according to BI-RADS, as well as the classification according to the matrix scheme, showed high inter- and intraobserver variability.

The sensitivity and specificity showed significant differences. For the classification with the matrix-scheme the results are 81% and 54%. The results of the classification according to BI-RADS are 86 % and 52 %. In the case of BI-RADS categories 2/3/4/5, malignant diagnoses were found in 13 %/24 %/ 53 %/86 % cases with the classification according BI-RADS. The classification with the matrix scheme found 8 %/32 %/48 %/84 % were malignant cases.

The matrix scheme shows an increase of malignant cases from diffuse to linear distribution and from typically benign to typically malignant morphology.

Discussion

The results of our study show that the complexity of microcalcification and the difficulties in detecting and describing them still remain the main problem when interpreting mammograms. The classification with the matrix scheme does not significantly reduce the observer variability neither does it improve the results of sensitivity and specificity. It is still a difficult question of how to link the descriptive findings to the assessment categories. Further evaluation regarding this task and the assignment of the matrix scheme will be necessary.

7. Literaturverzeichnis

1. Roque, A.C. and T.C. Andre, *Mammography and computerized decision systems: a review*. Ann N Y Acad Sci, 2002. **980**: p. 83-94.
2. Bauer, M., R. Schulz-Wendtland and D. von Fournier, [*Mass mammography screening. A means of reducing mortality in breast cancer?*]. Radiologe, 1988. **28**(3): p. 95-102.
3. Jemal, A., R. Siegel, E. Ward, T. Murray, J. Xu, C. Smigal and M.J. Thun, *Cancer statistics, 2006*. CA Cancer J Clin, 2006. **56**(2): p. 106-130.
4. Schön, D., J. Bertz, B. Görsch and, *Die Dachdokumentation Krebs. In: Gesundheitsbl - Gesundheitsforsch – Gesundheitsschutz 2004*. 2004. **47**: p. 429-436.
5. Caygill, C.P. and M.J. Hill, *Trends in European breast cancer incidence and possible etiology*. Tumori, 1991. **77**(2): p. 126-129.
6. Feuer, E.J., L.M. Wun, C.C. Boring, W.D. Flanders, M.J. Timmel and T. Tong, *The lifetime risk of developing breast cancer*. J Natl Cancer Inst, 1993. **85**(11): p. 892-897.
7. Landis, S.H., T. Murray, S. Bolden and P.A. Wingo, *Cancer statistics, 1998*. CA Cancer J Clin, 1998. **48**(1): p. 6-29.
8. Wingo, P.A., L.A. Ries, H.M. Rosenberg, D.S. Miller and B.K. Edwards, *Cancer incidence and mortality, 1973-1995: a report card for the U.S.* Cancer, 1998. **82**(6): p. 1197-1207.
9. *Krebs in Deutschland (2006), 5. überarbeitete, aktualisierte Ausgabe, Arbeitsgemeinschaft Bevölkerungsbezogener Krebsregister in Deutschland, Saarbrücken*. 2006.
10. Armstrong, K., A. Eisen and B. Weber, *Assessing the risk of breast cancer*. N Engl J Med, 2000. **342**(8): p. 564-571.
11. Geraedts, M. and C. Kugler, *Versorgungsforschung beim Mammakarzinom*. Gynäkologe, 2003. **36**: p. 878-883.
12. Chang-Claude, J. and, *Genetische Disposition bei Krebserkrankungen von Frauen, Schwerpunktthema FrauenLeben und Krebs*. Omnimed Verlagsgesellschaft, 1997. **8**: p. 595-597.
13. Jemal, A., R. Siegel, E. Ward, T. Murray, J. Xu and M.J. Thun, *Cancer statistics, 2007*. CA Cancer J Clin, 2007. **57**(1): p. 43-66.
14. HSK, *Onkologischer Schwerpunkt, Jahresbericht 2004, HSK Wiesbaden*. 2004.
15. Tabar, L. and P.B. Dean, *The control of breast cancer through mammography screening. What is the evidence?* Radiol Clin North Am, 1987. **25**(5): p. 993-1005.
16. Zuckerman, H.C., *The role of mammography in the diagnosis of breast cancer. In: Ariel IM, Cleary JB, eds. Breast cancer, diagnosis and treatment*. New York, NY: McGraw-Hill. 1987: p. 152-172.
17. Burstein, H.J., K. Polyak, J.S. Wong, S.C. Lester and C.M. Kaelin, *Ductal carcinoma in situ of the breast*. N Engl J Med, 2004. **350**(14): p. 1430-1441.

18. Tavassoli, F.A., *Breast pathology: rationale for adopting the ductal intraepithelial neoplasia (DIN) classification*. Nat Clin Pract Oncol, 2005. **2**(3): p. 116-117.
19. Viale, G., *Histopathology of primary breast cancer 2005*. Breast, 2005. **14**(6): p. 487-492.
20. Winchester, D.P. and E.A. Strom, *Standards for diagnosis and management of ductal carcinoma in situ (DCIS) of the breast*. American College of Radiology. American College of Surgeons. College of American Pathologists. Society of Surgical Oncology. CA Cancer J Clin, 1998. **48**(2): p. 108-128.
21. Ernster, V.L., R. Ballard-Barbash, W.E. Barlow, Y. Zheng, D.L. Weaver, G. Cutter, B.C. Yankaskas, R. Rosenberg, P.A. Carney, K. Kerlikowske, S.H. Taplin, N. Urban and B.M. Geller, *Detection of ductal carcinoma in situ in women undergoing screening mammography*. J Natl Cancer Inst, 2002. **94**(20): p. 1546-1554.
22. Kleinschmidt, O., *in Zweifel-Payer: Klinik der bösartigen Geschwulste*. 1927. **Bd IV, Leipzig; Hirzel 1927**.
23. Bird, R.E., T.W. Wallace and B.C. Yankaskas, *Analysis of cancers missed at screening mammography*. Radiology, 1992. **184**(3): p. 613-617.
24. van Dijck, J.A., A.L. Verbeek, J.H. Hendriks and R. Holland, *The current detectability of breast cancer in a mammographic screening program. A review of the previous mammograms of interval and screen-detected cancers*. Cancer, 1993. **72**(6): p. 1933-1938.
25. Goergen, S.K., J. Evans, G.P. Cohen and J.H. MacMillan, *Characteristics of breast carcinomas missed by screening radiologists*. Radiology, 1997. **204**(1): p. 131-135.
26. Mushlin, A.I., R.W. Kouides and D.E. Shapiro, *Estimating the accuracy of screening mammography: a meta-analysis*. Am J Prev Med, 1998. **14**(2): p. 143-153.
27. Brown, M.L., F. Houn, E.A. Sickles and L.G. Kessler, *Screening mammography in community practice: positive predictive value of abnormal findings and yield of follow-up diagnostic procedures*. AJR Am J Roentgenol, 1995. **165**(6): p. 1373-1377.
28. Kerlikowske, K., D. Grady, J. Barclay, S.D. Frankel, S.H. Ominsky, E.A. Sickles and V. Ernster, *Variability and accuracy in mammographic interpretation using the American College of Radiology Breast Imaging Reporting and Data System*. J Natl Cancer Inst, 1998. **90**(23): p. 1801-1809.
29. Humphrey, L.L., M. Helfand, B.K. Chan and S.H. Woolf, *Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force*. Ann Intern Med, 2002. **137**(5 Part 1): p. 347-360.
30. Tabar, L., S.W. Duffy, B. Vitak, H.H. Chen and T.C. Prevost, *The natural history of breast carcinoma: what have we learned from screening?* Cancer, 1999. **86**(3): p. 449-462.
31. Feig, S.A., *Decreased breast cancer mortality through mammographic screening: results of clinical trials*. Radiology, 1988. **167**(3): p. 659-665.
32. Duffy, S.W., L. Tabar, G. Fagerberg, A. Gad, O. Grontoft, M.C. South and N.E. Day, *Breast screening, prognostic factors and survival--results from the Swedish two county study*. Br J Cancer, 1991. **64**(6): p. 1133-1138.

33. Kerlikowske, K. and J. Barclay, *Outcomes of modern screening mammography*. J Natl Cancer Inst Monogr, 1997(22): p. 105-111.
34. Alexander, F.E., T.J. Anderson, H.K. Brown, A.P. Forrest, W. Hepburn, A.E. Kirkpatrick, B.B. Muir, R.J. Prescott and A. Smith, *14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening*. Lancet, 1999. **353**(9168): p. 1903-1908.
35. de Koning, H.J., *Assessment of nationwide cancer-screening programmes*. Lancet, 2000. **355**(9198): p. 80-81.
36. Sickles, E.A., *Findings at mammographic screening on only one standard projection: outcomes analysis*. Radiology, 1998. **208**(2): p. 471-475.
37. Hermann, G., C. Janus, I.S. Schwartz, A. Papatestas, D.G. Hermann and J.G. Rabinowitz, *Occult malignant breast lesions in 114 patients: relationship to age and the presence of microcalcifications*. Radiology, 1988. **169**(2): p. 321-324.
38. Murphy, W.A. and K. DeSchryver-Kecsckemeti, *Isolated clustered microcalcifications in the breast: radiologic-pathologic correlation*. Radiology, 1978. **127**(2): p. 335-341.
39. Sickles, E.A., *Breast calcifications: mammographic evaluation*. Radiology, 1986. **160**(2): p. 289-293.
40. Gilles, R., M. Meunier, O. Lucidarme, B. Zafrani, J.M. Guinebretiere, A.A. Tardivon, M. Le Gal, D. Vanel, S. Neuenschwander and R. Arriagada, *Clustered breast microcalcifications: evaluation by dynamic contrast-enhanced subtraction MRI*. J Comput Assist Tomogr, 1996. **20**(1): p. 9-14.
41. Sickles, E.A., *Mammographic features of "early" breast cancer*. AJR Am J Roentgenol, 1984. **143**(3): p. 461-464.
42. Kaiser, W.A., *[MR mammography]*. Radiologe, 1993. **33**(5): p. 292-299.
43. Langer-Cherbit, A., M. Le Gal, B. Asselain and S. Neuenschwander, *Breast cancer: zones of increased density mammographic features, correlated to clinical TNM and prognosis*. Eur J Radiol, 1997. **24**(1): p. 48-53.
44. Dengel, H., H. Leibersperger, E. Glatt, K. Reisner and W. Haase, *[Mammography after breast-preserving therapy of breast cancer]*. Rofo, 1987. **146**(5): p. 532-538.
45. Dershaw, D.D. and T.A. Chaglassian, *Mammography after prosthesis placement for augmentation or reconstructive mammoplasty*. Radiology, 1989. **170**(1 Pt 1): p. 69-74.
46. Paulus, D.D., H.I. Libshitz and E.D. Montague, *Malignant masses in the therapeutically irradiated breast*. AJR Am J Roentgenol, 1980. **135**(4): p. 789-795.
47. Sickles, E.A., R.A. Filly and P.W. Callen, *Breast cancer detection with sonography and mammography: comparison using state-of-the-art equipment*. AJR Am J Roentgenol, 1983. **140**(5): p. 843-845.
48. Elmore, J.G., C.K. Wells, C.H. Lee, D.H. Howard and A.R. Feinstein, *Variability in radiologists' interpretations of mammograms*. N Engl J Med, 1994. **331**(22): p. 1493-1499.
49. Sickles, E.A., D.E. Wolverton and K.E. Dee, *Performance parameters for screening and diagnostic mammography: specialist and general radiologists*. Radiology, 2002. **224**(3): p. 861-869.

50. Beam, C.A., E.F. Conant and E.A. Sickles, *Association of volume and volume-independent factors with accuracy in screening mammogram interpretation*. J Natl Cancer Inst, 2003. **95**(4): p. 282-290.
51. Anttinen, I., M. Pamilo, M. Soiva and M. Roiha, *Double reading of mammography screening films--one radiologist or two?* Clin Radiol, 1993. **48**(6): p. 414-421.
52. Thurfjell, E.L., K.A. Lernevall and A.A. Taube, *Benefit of independent double reading in a population-based mammography screening program*. Radiology, 1994. **191**(1): p. 241-244.
53. Shtern, F., *Digital mammography and related technologies: a perspective from the National Cancer Institute*. Radiology, 1992. **183**(3): p. 629-630.
54. Feig, S.A. and M.J. Yaffe, *Digital mammography, computer-aided diagnosis, and telemammography*. Radiol Clin North Am, 1995. **33**(6): p. 1205-1230.
55. Schmidt, R.A. and R.M. Nishikawa, *Clinical use of digital mammography: the present and the prospects*. J Digit Imaging, 1995. **8**(1 Suppl 1): p. 74-79.
56. Chan, H.P., M.A. Helvie, N. Petrick, B. Sahiner, D.D. Adler, C. Paramagul, M.A. Roubidoux, C.E. Blane, L.K. Joynt, T.E. Wilson, L.M. Hadjiiski and M.M. Goodsitt, *Digital mammography: observer performance study of the effects of pixel size on the characterization of malignant and benign microcalcifications*. Acad Radiol, 2001. **8**(6): p. 454-466.
57. Lewin, J.M., R.E. Hendrick, C.J. D'Orsi, P.K. Isaacs, L.J. Moss, A. Karellas, G.A. Sisney, C.C. Kuni and G.R. Cutter, *Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4,945 paired examinations*. Radiology, 2001. **218**(3): p. 873-880.
58. Venta, L.A., R.E. Hendrick, Y.T. Adler, P. DeLeon, P.M. Mengoni, A.M. Scharl, C.E. Comstock, L. Hansen, N. Kay, A. Coveler and G. Cutter, *Rates and causes of disagreement in interpretation of full-field digital mammography and film-screen mammography in a diagnostic setting*. AJR Am J Roentgenol, 2001. **176**(5): p. 1241-1248.
59. Fischmann, A., K.C. Siegmann, A. Wersbe, C.D. Claussen and M. Muller-Schimpfle, *Comparison of full-field digital mammography and film-screen mammography: image quality and lesion detection*. Br J Radiol, 2005. **78**(928): p. 312-315.
60. Fischer, U., F. Baum, S. Obenauer, S. Luftner-Nagel, D. von Heyden, R. Vosschenrich and E. Grabbe, *Comparative study in patients with microcalcifications: full-field digital mammography vs screen-film mammography*. Eur Radiol, 2002. **12**(11): p. 2679-2683.
61. Jung, H., *[Assessment of usefulness and risk of mammography screening with exclusive attention to radiation risk]*. Radiologe, 2001. **41**(4): p. 385-395.
62. Jung, H., *Is there a real risk of radiation-induced breast cancer for postmenopausal women?* Radiat Environ Biophys, 2001. **40**(2): p. 169-174.
63. Feig, S.A. and S.M. Ehrlich, *Estimation of radiation risk from screening mammography: recent trends and comparison with expected benefits*. Radiology, 1990. **174**(3 Pt 1): p. 638-647.
64. Jung, H., *[Mammography and radiation risk]*. Rofo, 1998. **169**(4): p. 336-343.

65. Heywang-Köbrunner, S.H., I. Schreer and R. Bässler, *Bildgebende Mammadiagnostik, Untersuchungstechnik, Befundmuster und Differentialdiagnostik in Mammographie, Sonographie und Kernspintomographie*. 1996: p. 221.
66. Yang, W.T. and G.M. Tse, *Sonographic, mammographic, and histopathologic correlation of symptomatic ductal carcinoma in situ*. AJR Am J Roentgenol, 2004. **182**(1): p. 101-110.
67. Schulz-Wendtland, R., E. Wenkel, U. Aichinger, M. Tartsch, I. Kuchar, A. Bodicker, C. Evertsz, H.O. Peitgen and W. Bautz, *[Film-screen mammography versus digital storage plate mammography: hard copy and monitor display of microcalcifications and focal findings--a retrospective clinical and histologic analysis]*. Rofo, 2003. **175**(9): p. 1220-1224.
68. Nishimura, S., K. Takahashi, N. Gomi, K. Tada, M. Makita, T. Tada, T. Iwase, M. Yoshimoto, F. Akiyama, G. Sakamoto and F. Kasumi, *What is the predictor for invasion in non-palpable breast cancer with microcalcifications?* Breast Cancer, 2004. **11**(1): p. 49-54.
69. Sun, Z., H.W. Liang and H.M. Xu, *Classification of breast microcalcifications: radiological-pathological correlation*. Chin Med J (Engl), 2005. **118**(17): p. 1429-1435.
70. (ACR), A.C.o.R., *American College of Radiology (ACR) (2003) Illustrated breast imaging reporting and data system (BI-RADS), 4th edn*. 2003.
71. Conway, B.J., J.L. McCrohan, F.G. Rueter and O.H. Suleiman, *Mammography in the eighties*. Radiology, 1990. **177**(2): p. 335-339.
72. McLelland, R., R.E. Hendrick, M.D. Zininger and P.A. Wilcox, *The American College of Radiology Mammography Accreditation Program*. AJR Am J Roentgenol, 1991. **157**(3): p. 473-479.
73. Obenauer, S., K.P. Hermann and E. Grabbe, *Applications and literature review of the BI-RADS classification*. Eur Radiol, 2005. **15**(5): p. 1027-1036.
74. Kopans, D.B., *Accuracy of mammographic interpretation*. N Engl J Med, 1994. **331**(22): p. 1521-1522.
75. Geller, B.M., W.E. Barlow, R. Ballard-Barbash, V.L. Ernster, B.C. Yankaskas, E.A. Sickles, P.A. Carney, M.B. Dignan, R.D. Rosenberg, N. Urban, Y. Zheng and S.H. Taplin, *Use of the American College of Radiology BI-RADS to report on the mammographic evaluation of women with signs and symptoms of breast disease*. Radiology, 2002. **222**(2): p. 536-542.
76. Sickles, E.A., *Probably benign breast lesions: when should follow-up be recommended and what is the optimal follow-up protocol?* Radiology, 1999. **213**(1): p. 11-14.
77. (ACR), A.C.o.R., *Illustrated Breast Imaging Reporting and Data System (BIRADS TM), 3rd edn*. 1998.
78. Orel, S.G., N. Kay, C. Reynolds and D.C. Sullivan, *BI-RADS categorization as a predictor of malignancy*. Radiology, 1999. **211**(3): p. 845-850.
79. Siegmann, K.C., A. Wersebe, A. Fischmann, N. Fersis, U. Vogel, C.D. Claussen and M. Muller-Schimpfle, *[Stereotactic vacuum-assisted breast biopsy--success, histologic accuracy, patient acceptance and optimizing the BI-RADSTM-correlated indication]*. Rofo, 2003. **175**(1): p. 99-104.

80. Sickles, E.A., *Periodic mammographic follow-up of probably benign lesions: results in 3,184 consecutive cases*. Radiology, 1991. **179**(2): p. 463-468.
81. Muller-Schimpfle, M., A. Wersebe, T. Xydeas, A. Fischmann, U. Vogel, N. Fersis, C.D. Claussen and K. Siegmann, *Microcalcifications of the breast: how does radiologic classification correlate with histology?* Acta Radiol, 2005. **46**(8): p. 774-781.
82. Kerlikowske, K., D. Grady, J. Barclay, E.A. Sickles and V. Ernster, *Likelihood ratios for modern screening mammography. Risk of breast cancer based on age and mammographic interpretation*. Jama, 1996. **276**(1): p. 39-43.
83. Saarenmaa, I., T. Salminen, U. Geiger, P. Heikkinen, S. Hyvarinen, J. Isola, V. Kataja, M.L. Kokko, R. Kokko, E. Kumpulainen, A. Karkkainen, J. Pakkanen, P. Peltonen, A. Piironen, A. Salo, M.L. Talviala and M. Haka, *The effect of age and density of the breast on the sensitivity of breast cancer diagnostic by mammography and ultasonography*. Breast Cancer Res Treat, 2001. **67**(2): p. 117-123.
84. Beam, C.A., P.M. Layde and D.C. Sullivan, *Variability in the interpretation of screening mammograms by US radiologists. Findings from a national sample*. Arch Intern Med, 1996. **156**(2): p. 209-213.
85. (EUREF) and E.P.f.t.Q.C.o.t.P.a.T.A.o.M. Screening., *“European Guidelines for Quality Assurance in Mammography Screening” Appendix 5: Digital mammography*. 1999.
86. Berg, W.A., C. Campassi, P. Langenberg and M.J. Sexton, *Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment*. AJR Am J Roentgenol, 2000. **174**(6): p. 1769-1777.
87. de Lafontan, B., J.P. Daures, B. Salicru, F. Eynius, J. Mihura, P. Rouanet, J.L. Lamarque, A. Naja and H. Pujol, *Isolated clustered microcalcifications: diagnostic value of mammography--series of 400 cases with surgical verification*. Radiology, 1994. **190**(2): p. 479-483.
88. Fondrinier, E., G. Lorimier, V. Guerin-Boblet, A.F. Bertrand, C. Mayras and N. Dauver, *Breast microcalcifications: multivariate analysis of radiologic and clinical factors for carcinoma*. World J Surg, 2002. **26**(3): p. 290-296.
89. Le Gal, M., J.C. Durand, M. Laurent and D. Pellier, *[Management following mammography revealing grouped microcalcifications without palpable tumor]*. Nouv Presse Med, 1976. **5**(26): p. 1623-1627.
90. Leborgne, R., *[Diagnosis of tumors of the breast by simple radiography.]*. Mem Acad Chir (Paris), 1951. **77**(14-15): p. 449-454.
91. Liberman, L., A.F. Abramson, F.B. Squires, J.R. Glassman, E.A. Morris and D.D. Dershaw, *The breast imaging reporting and data system: positive predictive value of mammographic features and final assessment categories*. AJR Am J Roentgenol, 1998. **171**(1): p. 35-40.
92. Bock, K., J. Iwinska-Zelder, V.F. Duda, C. Bonwetsch, G. Rode, P. Hadji, K.J. Klose and K.D. Schulz, *[Validity of the breast imaging reporting and data system BI-RADS(TM) for clinical mammography in men]*. Rofo, 2001. **173**(11): p. 1019-1024.
93. Boehm, T., E. Garzoli and B. Marincek, *[Differential diagnosis of benign and malignant mammary lesions with special regard to the BI-RADS((R)) classification system in mammography]*. Gynakol Geburtshilfliche Rundsch, 2002. **42**(4): p. 191-200.

94. Chabriaux, J., K. Kinkel and F. Thibault, *[Using BI-RADS in French]*. J Radiol, 2002. **83**(4 Pt 2): p. 531-534.
95. Gulsun, M., F.B. Demirkazik and M. Ariyurek, *Evaluation of breast microcalcifications according to Breast Imaging Reporting and Data System criteria and Le Gal's classification*. Eur J Radiol, 2003. **47**(3): p. 227-231.
96. Obenauer, S., U. Fischer, F. Baum, S. Dammert, L. Fuzesi and E. Grabbe, *[Stereotactic vacuum core biopsy of clustered microcalcifications classified as B1-RADS 3]*. Rofo, 2001. **173**(8): p. 696-701.
97. Baker, J.A., P.J. Kornguth and C.E. Floyd, Jr., *Breast imaging reporting and data system standardized mammography lexicon: observer variability in lesion description*. AJR Am J Roentgenol, 1996. **166**(4): p. 773-778.
98. Berube, M., B. Curpen, P. Ugolini, L. Lalonde and D. Ouimet-Oliva, *Level of suspicion of a mammographic lesion: use of features defined by BI-RADS lexicon and correlation with large-core breast biopsy*. Can Assoc Radiol J, 1998. **49**(4): p. 223-228.
99. Lacquement, M.A., D. Mitchell and A.B. Hollingsworth, *Positive predictive value of the Breast Imaging Reporting and Data System*. J Am Coll Surg, 1999. **189**(1): p. 34-40.
100. Zonderland, H.M., T.L. Pope, Jr. and A.J. Nieborg, *The positive predictive value of the breast imaging reporting and data system (BI-RADS) as a method of quality assessment in breast imaging in a hospital population*. Eur Radiol, 2004. **14**(10): p. 1743-1750.
101. Berg, W.A., C.J. D'Orsi, V.P. Jackson, L.W. Bassett, C.A. Beam, R.S. Lewis and P.E. Crewson, *Does training in the Breast Imaging Reporting and Data System (BI-RADS) improve biopsy recommendations or feature analysis agreement with experienced breast imagers at mammography?* Radiology, 2002. **224**(3): p. 871-880.
102. Kim, S.J., E.A. Morris, L. Liberman, D.J. Ballon, L.R. La Trenta, O. Hadar, A. Abramson and D.D. Dershaw, *Observer variability and applicability of BI-RADS terminology for breast MR imaging: invasive carcinomas as focal masses*. AJR Am J Roentgenol, 2001. **177**(3): p. 551-557.
103. Lehman, C., S. Holt, S. Peacock, E. White and N. Urban, *Use of the American College of Radiology BI-RADS guidelines by community radiologists: concordance of assessments and recommendations assigned to screening mammograms*. AJR Am J Roentgenol, 2002. **179**(1): p. 15-20.
104. Liberman, L. and J.H. Menell, *Breast imaging reporting and data system (BI-RADS)*. Radiol Clin North Am, 2002. **40**(3): p. 409-430, v.
105. Vitiello SM, P.L., Tocino I, and L.C. Horvarth LJ, *Understanding of BI-RADS among referring clinicians: do they get it? (abstract)*. J Roentgenol, 2000. **174**(Suppl)(64).
106. Pijnappel, R.M., P.H. Peeters, J.H. Hendriks and W.P. Mali, *Reproducibility of mammographic classifications for non-palpable suspect lesions with microcalcifications*. Br J Radiol, 2004. **77**(916): p. 312-314.
107. Dinkel, H.P., A.M. Gassel and A. Tschammler, *Is the appearance of microcalcifications on mammography useful in predicting histological grade of malignancy in ductal cancer in situ?* Br J Radiol, 2000. **73**(873): p. 938-944.
108. Yunus, M., N. Ahmed, I. Masroor and J. Yaqoob, *Mammographic criteria for determining the diagnostic value of microcalcifications in the detection of early breast cancer*. J Pak

- Med Assoc, 2004. **54**(1): p. 24-29.
109. Sakka, E., A. Prentza and D. Koutsouris, *Classification algorithms for microcalcifications in mammograms (Review)*. Oncol Rep, 2006. **15 Spec no.:** p. 1049-1055.
110. Hall, F.M., J.M. Storella, D.Z. Silverstone and G. Wyshak, *Nonpalpable breast lesions: recommendations for biopsy based on suspicion of carcinoma at mammography*. Radiology, 1988. **167**(2): p. 353-358.
111. Stomper, P.C., J.L. Connolly, J.E. Meyer and J.R. Harris, *Clinically occult ductal carcinoma in situ detected with mammography: analysis of 100 cases with radiologic-pathologic correlation*. Radiology, 1989. **172**(1): p. 235-241.
112. Stomper, P.C., *Mammographic local staging of breast carcinoma: methods and limitations*. Breast Dis, 2001. **13**: p. 85-95.
113. Darling, M.L., D.N. Smith, S.C. Lester, C. Kaelin, D.L. Selland, C.M. Denison, P.J. DiPiro, D.I. Rose, E. Rhei and J.E. Meyer, *Atypical ductal hyperplasia and ductal carcinoma in situ as revealed by large-core needle breast biopsy: results of surgical excision*. AJR Am J Roentgenol, 2000. **175**(5): p. 1341-1346.
114. Ciatto, S., L. Cataliotti and V. Distante, *Nonpalpable lesions detected with mammography: review of 512 consecutive cases*. Radiology, 1987. **165**(1): p. 99-102.
115. Egan, R.L., M.B. McSweeney and C.W. Sewell, *Intramammary calcifications without an associated mass in benign and malignant diseases*. Radiology, 1980. **137**(1 Pt 1): p. 1-7.
116. Homer, M.J., *Breast imaging: pitfalls, controversies, and some practical thoughts*. Radiol Clin North Am, 1985. **23**(3): p. 459-472.
117. Bassett, L.W., *Mammographic analysis of calcifications*. Radiol Clin North Am, 1992. **30**(1): p. 93-105.
118. Sickles, E.A., *Mammographic features of 300 consecutive nonpalpable breast cancers*. AJR Am J Roentgenol, 1986. **146**(4): p. 661-663.
119. Cady, B. and M.A. Chung, *The prevention of invasive breast carcinoma*. Cancer, 2004. **101**(10): p. 2147-2151.
120. Stomper, P.C. and J.L. Connolly, *Ductal carcinoma in situ of the breast: correlation between mammographic calcification and tumor subtype*. AJR Am J Roentgenol, 1992. **159**(3): p. 483-485.
121. Bonnett, M., T. Wallis, M. Rossmann, N.L. Pernick, K.A. Carolin, M. Segel, D. Bouwman and D. Visscher, *Histologic and radiographic analysis of ductal carcinoma in situ diagnosed using stereotactic incisional core breast biopsy*. Mod Pathol, 2002. **15**(2): p. 95-101.
122. Krecke, K.N. and J.J. Gisvold, *Invasive lobular carcinoma of the breast: mammographic findings and extent of disease at diagnosis in 184 patients*. AJR Am J Roentgenol, 1993. **161**(5): p. 957-960.
123. Hilleren, D.J., I.T. Andersson, K. Lindholm and F.S. Linnell, *Invasive lobular carcinoma: mammographic findings in a 10-year experience*. Radiology, 1991. **178**(1): p. 149-154.
124. Mendelson, E.B., K.M. Harris, N. Doshi and H. Tobon, *Infiltrating lobular carcinoma:*

- mammographic patterns with pathologic correlation. AJR Am J Roentgenol*, 1989. **153**(2): p. 265-271.
125. Le Gal, M., L. Ollivier, B. Asselain, M. Meunier, M. Laurent, P. Vielh and S. Neuenschwander, *Mammographic features of 455 invasive lobular carcinomas. Radiology*, 1992. **185**(3): p. 705-708.
126. Ashikari, R., A.G. Huvos, J.A. Urban and G.F. Robbins, *Infiltrating lobular carcinoma of the breast. Cancer*, 1973. **31**(1): p. 110-116.
127. Feig, S.A., G.S. Shaber, A. Patchefsky, G.F. Schwartz, J. Edeiken, H.I. Libshitz, R. Nerlinger, R.F. Curley and J.D. Wallace, *Analysis of clinically occult and mammographically occult breast tumors. AJR Am J Roentgenol*, 1977. **128**(3): p. 403-408.
128. Moskowitz, M., *The predictive value of certain mammographic signs in screening for breast cancer. Cancer*, 1983. **51**(6): p. 1007-1011.
129. Hall, F.M., *Malignancy in BI-RADS category 3 mammographic lesions. Radiology*, 2002. **225**(3): p. 918-919; author reply 919-920.
130. Varas, X., F. Leborgne and J.H. Leborgne, *Nonpalpable, probably benign lesions: role of follow-up mammography. Radiology*, 1992. **184**(2): p. 409-414.
131. Rosen, E.L., J.A. Baker and M.S. Soo, *Malignant lesions initially subjected to short-term mammographic follow-up. Radiology*, 2002. **223**(1): p. 221-228.
132. Lev-Toaff, A.S., S.A. Feig, V.L. Saitas, G.C. Finkel and G.F. Schwartz, *Stability of malignant breast microcalcifications. Radiology*, 1994. **192**(1): p. 153-156.
133. Hadjiiski, L., H.P. Chan, B. Sahiner, M.A. Helvie, M.A. Roubidoux, C. Blane, C. Paramagul, N. Petrick, J. Bailey, K. Klein, M. Foster, S. Patterson, D. Adler, A. Nees and J. Shen, *Improvement in radiologists' characterization of malignant and benign breast masses on serial mammograms with computer-aided diagnosis: an ROC study. Radiology*, 2004. **233**(1): p. 255-265.
134. Monsees, B.S., *Evaluation of breast microcalcifications. Radiol Clin North Am*, 1995. **33**(6): p. 1109-1121.
135. Burnside, E.S., J.E. Ochsner, K.J. Fowler, J.P. Fine, L.R. Salkowski, D.L. Rubin and G.A. Sisney, *Use of microcalcification descriptors in BI-RADS 4th edition to stratify risk of malignancy. Radiology*, 2007. **242**(2): p. 388-395.
136. Varas, X., J.H. Leborgne, F. Leborgne, J. Mezzera, S. Jaumandreu and F. Leborgne, *Revisiting the mammographic follow-up of BI-RADS category 3 lesions. AJR Am J Roentgenol*, 2002. **179**(3): p. 691-695.
137. Mendez, A., F. Cabanillas, M. Echenique, K. Malekshamran, I. Perez and E. Ramos, *Evaluation of Breast Imaging Reporting and Data System Category 3 mammograms and the use of stereotactic vacuum-assisted breast biopsy in a nonacademic community practice. Cancer*, 2004. **100**(4): p. 710-714.
138. Franceschi, D., J. Crowe, R. Zollinger, R. Duchesneau, R. Shenk, G. Stefanek and J.M. Shuck, *Biopsy of the breast for mammographically detected lesions. Surg Gynecol Obstet*, 1990. **171**(6): p. 449-455.
139. Le Gal, M., G. Chavanne and D. Pellier, *[Diagnostic value of clustered*

- microcalcifications discovered by mammography (apropos of 227 cases with histological verification and without a palpable breast tumor)]. Bull Cancer, 1984. 71(1): p. 57-64.*
140. Bocker, W., T. Decker, M. Ruhnke and W. Schneider, [*Ductal hyperplasia and ductal carcinoma in situ. Definition--classification--differential diagnosis*]. Pathologe, 1997. **18**(1): p. 3-18.
141. Leichter, I., S. Fields, R. Nirel, P. Bamberger, B. Novak, R. Lederman and S. Buchbinder, *Improved mammographic interpretation of masses using computer-aided diagnosis*. Eur Radiol, 2000. **10**(2): p. 377-383.
142. Chan, H.P., B. Sahiner, M.A. Helvie, N. Petrick, M.A. Roubidoux, T.E. Wilson, D.D. Adler, C. Paramagul, J.S. Newman and S. Sanjay-Gopal, *Improvement of radiologists' characterization of mammographic masses by using computer-aided diagnosis: an ROC study*. Radiology, 1999. **212**(3): p. 817-827.
143. Jiang, Y., R.M. Nishikawa, R.A. Schmidt, A.Y. Toledano and K. Doi, *Potential of computer-aided diagnosis to reduce variability in radiologists' interpretations of mammograms depicting microcalcifications*. Radiology, 2001. **220**(3): p. 787-794.
144. Doi, K., M.L. Giger, R.M. Nishikawa and R.A. Schmidt, *Computer-Aided Diagnosis of Breast Cancer on Mammograms*. Breast Cancer, 1997. **4**(4): p. 228-233.
145. Warren Burhenne, L.J., S.A. Wood, C.J. D'Orsi, S.A. Feig, D.B. Kopans, K.F. O'Shaughnessy, E.A. Sickles, L. Tabar, C.J. Vyborny and R.A. Castellino, *Potential contribution of computer-aided detection to the sensitivity of screening mammography*. Radiology, 2000. **215**(2): p. 554-562.
146. Veldkamp, W.J., N. Karssemeijer, J.D. Otten and J.H. Hendriks, *Automated classification of clustered microcalcifications into malignant and benign types*. Med Phys, 2000. **27**(11): p. 2600-2608.
147. Chan, H.P., S.C. Lo, B. Sahiner, K.L. Lam and M.A. Helvie, *Computer-aided detection of mammographic microcalcifications: pattern recognition with an artificial neural network*. Med Phys, 1995. **22**(10): p. 1555-1567.
148. Schmidt, F., E. Sorantin, C. Szepesvari, E. Graif, M. Becker, H. Mayer and K. Hartwagner, *An automatic method for the identification and interpretation of clustered microcalcifications in mammograms*. Phys Med Biol, 1999. **44**(5): p. 1231-1243.
149. Boccignone, G., A. Chianese and A. Picariello, *Computer aided detection of microcalcifications in digital mammograms*. Comput Biol Med, 2000. **30**(5): p. 267-286.
150. Markopoulos, C., E. Kouskos, K. Koufopoulos, V. Kyriakou and J. Gogas, *Use of artificial neural networks (computer analysis) in the diagnosis of microcalcifications on mammography*. Eur J Radiol, 2001. **39**(1): p. 60-65.
151. Freer, T.W. and M.J. Ulissey, *Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center*. Radiology, 2001. **220**(3): p. 781-786.
152. Getty, D.J., R.M. Pickett, C.J. D'Orsi and J.A. Swets, *Enhanced interpretation of diagnostic images*. Invest Radiol, 1988. **23**(4): p. 240-252.
153. Doi, K., H. MacMahon, S. Katsuragawa, R.M. Nishikawa and Y. Jiang, *Computer-aided diagnosis in radiology: potential and pitfalls*. Eur J Radiol, 1999. **31**(2): p. 97-109.

154. Birdwell, R.L., D.M. Ikeda, K.F. O'Shaughnessy and E.A. Sickles, *Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection*. Radiology, 2001. **219**(1): p. 192-202.
155. Jiang, Y., R.M. Nishikawa, R.A. Schmidt, C.E. Metz, M.L. Giger and K. Doi, *Improving breast cancer diagnosis with computer-aided diagnosis*. Acad Radiol, 1999. **6**(1): p. 22-33.
156. Lederman, R., I. Leichter, S. Buchbinder, B. Novak, P. Bamberger and S. Fields, *Stratification of mammographic computerized analysis by BI-RADS categories*. Eur Radiol, 2003. **13**(2): p. 347-353.
157. Baker, J.A., P.J. Kornguth, J.Y. Lo, M.E. Williford and C.E. Floyd, Jr., *Breast cancer: prediction with artificial neural network based on BI-RADS standardized lexicon*. Radiology, 1995. **196**(3): p. 817-822.
158. Leichter, I., R. Lederman, P. Bamberger, B. Novak, S. Fields and S.S. Buchbinder, *The use of an interactive software program for quantitative characterization of microcalcifications on digitized film-screen mammograms*. Invest Radiol, 1999. **34**(6): p. 394-400.
159. Buchbinder, S.S., I.S. Leichter, R.B. Lederman, B. Novak, P.N. Bamberger, M. Sklair-Levy, G. Yarmish and S.I. Fields, *Computer-aided classification of BI-RADS category 3 breast lesions*. Radiology, 2004. **230**(3): p. 820-823.
160. Gur, D., J.H. Sumkin, H.E. Rockette, M. Ganott, C. Hakim, L. Hardesty, W.R. Poller, R. Shah and L. Wallace, *Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system*. J Natl Cancer Inst, 2004. **96**(3): p. 185-190.
161. Astley, S.M., *Computer-aided detection for screening mammography*. Acad Radiol, 2004. **11**(10): p. 1139-1143.
162. Baker, J.A., J.Y. Lo, D.M. DeLong and C.E. Floyd, *Computer-aided detection in screening mammography: variability in cues*. Radiology, 2004. **233**(2): p. 411-417.

8. Danksagung

An dieser Stelle gilt Herrn Professor Dr. S. Miller mein besonderer Dank für die Überlassung dieses Promotionsthemas und die Vertretung vor der Fakultät.

Mein spezieller Dank gilt Herrn Dr. A. Fischmann für das Ermöglichen, Betreuen und Leiten dieser Arbeit. Sein Interesse an diesem Projekt und sein Engagement haben das Gelingen erst möglich gemacht.

Vielen Dank für die Unterstützung bei der Bearbeitung und Erhebung der Daten, die Hilfsbereitschaft bei allen aufkommenden Fragen und Problemen, die Durchsicht und Korrektur dieser Arbeit.

Des Weiteren möchte ich mich sehr herzlich bei Frau Dr. K. Siegmann, Frau Dr. A. Wersebe und Frau Dr. T. Xydeas für Ihre engagierte Mitarbeit an dieser Arbeit bedanken.

Frau Dr. T. Herberts danke ich für die Statistische Auswertung der Daten und Beratung bei der Beurteilung der Ergebnisse.

Für das Ermöglichen des Studiums, der Promotion und der Unterstützung während der letzten Jahre möchte ich mich bei meiner Familie bedanken. Für seine liebevolle und geduldige Unterstützung danke ich Thomas von ganzem Herzen.

9. Lebenslauf

Persönliche Daten

Name	Rebecca Sarah Läufe
Geburtsdatum	15. Dezember 1980
Geburtsort	Radolfzell am Bodensee
Eltern	Wolfgang Läufe, Dr. Rita Läufe, geb. Löffler
Geschwister	Simone Läufe
Familienstand	ledig
Staatsangehörigkeit	Deutsch

Schulbildung

1987 -1991	Grundschule Volkertshausen
1991 -2000	Hegau-Gymnasium Singen/Htwl.
Schulabschluss	Allgemeine Hochschulreife, Note: 1,6

Hochschulausbildung

04/2001	Beginn des Studiums der Humanmedizin an der Eberhard-Karls-Universität Tübingen
04/2003	Ärztliche Vorprüfung - Physikum (ÄAppO –ALT-) Note: 2,0
04/2004	Beginn der Promotion an der Radiologischen Universitätsklinik Tübingen.
02/2006	Beginn des Praktischen Jahres am Klinikum Esslingen, Lehrkrankenhaus der Universität Tübingen Wahlfach: Gynäkologie und Geburtshilfe
22.11/2007	Zweiter Abschnitt der ärztlichen Prüfung (ÄAppO 2002 –NEU-) Note: 2,0
12/2007	Approbation
04/2008	Assistenzärztin im Klinikum Karlsruhe-Frauenklinik